Author Manuscript Author Manuscript Author Manuscript Author Manuscript

HHS Public Access
Author manuscript
Lang Learn Dev. Author manuscript; available in PMC 2023 February 23.

Published in final edited form as:
Lang Learn Dev. 2023 ; 19(1): 95–123. doi:10.1080/15475441.2022.2047689.

Number Stroop Effects in Arabic Digits and ASL Number Signs: The Impact of Age and Setting of Language Acquisition

Nina Semushina,
Rachel Mayberry
Department of Linguistics, La Jolla, UCSD, San Diego, California, USA

Abstract

Multiple studies have reported mathematics underachievement for students who are deaf, but the onset, scope, and causes of this phenomenon remain understudied. Early language deprivation might be one factor influencing the acquisition of numbers. In this study, we investigated a basic and fundamental mathematical skill, automatic magnitude processing, in two formats (Arabic digits and American Sign Language number signs) and the influence of age of first language exposure on both formats by using two versions of the Number Stroop Test. We compared the performance of individuals born deaf who experienced early language deprivation to that of individuals born deaf who experienced sign language in early life and hearing second language learners of ASL. In both formats of magnitude representation, late first language learners demonstrated overall slower reaction times. They were also less accurate on incongruent trials but performed no differently from early signers and second language learners on other trials. When magnitude was represented by Arabic digits, late first language learners exhibited robust Number Stroop Effects, suggesting automatic magnitude processing, but they also demonstrated a large speed difference between size and number judgments not observed in the other groups. In a task with ASL number signs, the Number Stroop Effect was not found in any group, suggesting that magnitude representation might be format-specific, in line with the results from several other languages. Late first language learners also demonstrate unusual patterns of slower reaction time for neutral rather than incongruent stimuli. Together, the results show that early language deprivation affects the ability to automatically judge quantities expressed both linguistically and by Arabic digits, but that it can be acquired later in life when language is available. Contrary to previous studies that find differences in speed of number processing between deaf and hearing participants, we find that when language is acquired early in life, deaf signers perform identically to hearing participants.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

CONTACT Nina Semushina, nsemushina@uchicago.edu, Department of Linguistics, La Jolla, UCSD, San Diego, CA 92093-0021, USA.

Disclosure statement
No potential conflict of interest was reported by the authors.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Introduction

Mathematic underachievement and language deprivation

Several studies have reported delays in number acquisition and mathematical development in deaf students. The delays are often attributed to hearing loss, unrelated to setting of education (Kritzer, 2009; Traxler, 2000; D. Wood et al., 1983; H. A. Wood et al., 1984), and are hypothesized to persist into adulthood, since deaf college students in several experiments processed magnitude more slowly than hearing students (Bull et al., 2005; Epstein et al., 1994). These delays are often found in studies that use standardized school tests with spoken language (Gottardis et al., 2011). However, other studies have not identified such delays in children or adults who are deaf, especially when looking at individual aspects of mathematical development (Bull et al., 2006; Gottardis et al., 2011; Iversen et al., 2004). For example, deaf preschoolers outperformed their hearing peers on some spatial and temporal numerical tasks (Arfé et al., 2011; Zarfaty et al., 2004), which indicates that hearing loss per se does not impact quantity discrimination and number reasoning at young ages.

Proficiency in sign language positively correlates with mathematical achievement in deaf children (Henner, Pagliaro, Sullivan, & Hoffmeister, 2021; Hrastinski & Wilbur, 2016). Moreover, a positive impact of bimodal bilingual education on school performance has been demonstrated for deaf children with various language backgrounds: in mathematics specifically (Lange et al., 2013) and in other aspects such as reading and spoken language proficiency (Henner et al., 2015; Hermans et al., 2008. Deaf children from deaf families who have access to sign language at home show an advantage in standardized mathematic assessments, scoring on par or even better than their hearing peers (Henner, Pagliaro, Sullivan, & Hoffmeister, i2021).

However, deaf and hard of hearing students do not constitute a homogenous group, but vary in life experience and cultural and language background. Fewer than 10% of deaf children are born into deaf families using sign language; the remaining 90% receive limited or no sign language (Mitchell & Karchmer, 2004) and thus many experience reduced language exposure early in life.

Lack of language exposure in early life limits early number exposure (such as number words or signs, grammatical plural markers, and the context for numbers in reading and storytelling) that are foundational (Anderson et al., 2005; Kritzer, 2009; Pagliaro & Kritzer, 2013). It also negatively affects working memory (Marshall et al., 2015), which is necessary for successful acquisition of numbers and mathematics (Holmes & Adams, 2006). It has been shown that the working memory of deaf children from deaf families (6–11 years old) who had early access to sign language is not different from that of hearing controls on non-verbal working memory assessments; whereas deaf children with later language access scored significantly lower (Marshall et al., 2015). In addition, not all deaf individuals have access to a natural sign language, even by school age, experiencing severe language deprivation: a biological state interfering with the development and maturation of the brain neurolinguistic structures (Cheng et al., 2019; Humphries et al., 2016; Pénicaud et al., 2013). In most severe cases, the first sign language input is received only post-childhood, past the sensitive period for language acquisition (see R. I Mayberry & Kluender, 2018 for detailed

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

discussion of language deprivation and sensitive period). The effect of language deprivation on number acquisition is understudied. Here, we report one approach to investigating several unanswered questions that can illuminate mathematical development in this population.

First, we ask if early language deprivation affects automatic number processing, a basic skill that is needed for calculation. To answer the question, we compared performance on a Number Stroop Test with Arabic digits of late first language learners of ASL with two control groups, deaf early childhood learners of American Sign Language (ASL) and hearing second language learners of ASL. Second, there are many ways to represent number symbolically. At an initial stage, number acquisition involves the interaction of different types of representation: number signs and digits. This fact requires that we ask whether the automatic processing of magnitudes depends upon the format or alternatively is similar across digits and number words. The results have been controversial (see Cohen Kadosh and Walsh (2009) for a review). Therefore, we asked participants to do a Number Stroop Task with ASL number signs to determine whether the Number Stroop Effect typically found for Arabic digits is also evoked by number signs.

The third question we asked was whether language deprivation affects magnitude processing for both number formats similarly, taking into account that late first language learners might have been exposed to Arabic digits earlier than to number signs. The relation between acquisition and the processing of digits and linguistic numbers in children is difficult to disentangle due to their relatively simultaneous exposure to both number formats. Overall, the relationship between language and number remains a topic of considerable debate (Carey, 2009; Gelman & Butterworth, 2005; Spelke, 2017). Research with individuals who acquired number signs and Arabic digits on different developmental timelines can contribute to our understanding of this relationship.

We begin by reviewing the literature on language deprivation, the Number Stroop Task in digits and number words/signs followed by a description of the current study. The methods, results, and summaries of Arabic Digit and ASL tasks are presented separately, followed by brief discussions, and followed by the general discussion.

Language deprivation: impact on language and number development

Late first language learners are congenitally deaf individuals who did not have early access to natural sign language and/or early spoken language intervention and thus began first language acquisition around or post puberty. These individuals do not demonstrate cognitive impairments. They were not socially deprived, unlike cases of isolated children (Fromkin et al., 1974; Koluchová, 1972, 1976). Some late first language learners develop homesigns – gestural communicative systems used with their families before they begin learning their first language later in life. However, delayed exposure to the first language has long-lasting detrimental effect on language proficiency and language outcome in comparison to both first and second language learners (Cheng & Mayberry, 2019, 2021; Ferjan Ramirez et al., 2016; R.I. Mayberry & Lock, 2003). In late first language learners, years of experience do not predict language proficiency: if language acquisition has started late, native-like proficiency is not achieved even after considerable exposure to language, suggesting an effect of a sensitive period. It has been shown that initially the language acquisition progress

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

of late first language learners follows the same milestones as children learning language with respect to the acquisition of vocabulary and word combinations (Berk & Lillo-Martin, 2012; Ferjan Ramirez, Lieberman, et al., 2013). Late learners are able to successfully master some mono-clausal, but not more syntactically complex syntactic structures (Boudreault & Mayberry, 2006; Cheng & Mayberry, 2019; Fromkin et al., 1974; Mayberry, Cheng, Hatrak, & Ilkbasaran, 2017; Mayberry, Davenport, Roth, & Halgren, 2018; Newport, 1990). However, there have not been systematic studies of the effect of severe language deprivation on number reasoning.

Being immersed in a numerate society, late first language learners often learn Arabic digits earlier than they acquire language and conventional number signs. Work with deaf Nicaraguans (Flaherty & Senghas, 2011) showed that one of the participants who lacked early access to language was able to produce and interpret large numbers written with Arabic digits, but was not able to recite a counting list in Nicaraguan Sign Language. While performing well on matching tasks with stimuli physically present (i.e., when the participants had to match the number of items that the experimenter physically presented to them in real time), this participant did not perform well on an ephemeral matching task (when the items that the participants had to match were no longer physically present after they were presented). Thus, Flaherty and Senghas (2011) concluded that knowledge of Arabic digits alone is insufficient for successful mental tracking of quantities. At the same time, by testing a diverse group of subjects with various backgrounds, they also showed that when a language is finally available, the counting sequence can be learned in adulthood. However, the effects of severe language deprivation on number processing is unknown.

Number Stroop Test: Arabic digits and number signs

Automatic magnitude processing is a basic skill that implies understanding of magnitude and is necessary for skilled calculation. It has been extensively studied with the Number Stroop Test Paradigm (Algom et al., 1996; Besner & Coltheart, 1979; Bull et al., 2006; Gebuis et al., 2009; Henik & Tzelgov, 1982; Kaufmann et al., 2008; Liu et al., 2006; Pansky & Algom, 2002; Razpurker-Apfeld & Koriat, 2006; Schwarz & Heinze, 1998). During the test, participants compare pairs of stimuli that differ both in physical size and magnitude and are instructed to choose the stimulus that is “bigger,” but the task focuses only on one aspect of the stimuli (size or magnitude). The stimuli vary in congruity (as illustrated by Figure 1): in congruent trials, size, and number information align (3 is smaller than 5). In incongruent trials, size information contradicts the numerical dimension (3 is physically larger than 5), and in neutral trials the digits differ only in a relevant dimension (size or magnitude). Reaction time (RT) across studies shows a facilitation effect (RT in congruent trials is faster than in neutral trials), as well as interference effects (RT in incongruent trials is slower than neutral). While interference is present in both numerical and size comparisons, facilitation may be absent in size comparison, since neutral stimuli may be particularly easy to process, with less chances to further speed up the processing in congruent trials (Girelli et al., 2000; Henik & Tzelgov, 1982)

This size congruity effect, or Number Stroop Effect, has been interpreted as evidence in favor of automatic parallel processing of both magnitude and size information: irrelevant

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

information was accessed even in the trials where it was not beneficial. The Number Stroop effect emerges in children after the start of schooling (Girelli et al., 2000; Rubinsten et al., 2002; White et al., 2012) and has been studied to assess automatic magnitude processing in children with varying degrees of mathematical achievement (Heine et al., 2010) or mathematical disabilities (Ashkenazi et al., 2009; Rousselle & Noël, 2007; Rubinsten & Henik, 2005). To date, automatic magnitude processing has not yet been studied in adults who learned their first language late in life.

Importantly, magnitudes can be expressed symbolically not only through conventional mathematic symbols as described above, but also linguistically through numerals. Studies with Arabic digits unambiguously suggest automaticity of unintentional number processing, but when numbers are represented linguistically, the results show great variability. The presence of the Number Stroop Effect when participants are reading number words appears to be specific to a language, or even a particular writing system. In Japanese, it has been found only in ideographic Kanji script, but not the syllabic Kana script (Takahashi & Green, 1983). In Hebrew, Number Stroop Effect with the gematric numerals, which are letters of the alphabet that stand for numbers, was similar to the one with Arabic digits (Razpurker-Apfeld & Koriat, 2006), but the effect with Hebrew number words was not (Cohen Kadosh et al., 2008). While the first linguistic Number Stroop Effect study did not find the effect in English (Besner & Coltheart, 1979), later Vaid (1985) found such an effect, and hypothesized that size congruity in number words may be language-specific, such that its processing depends upon the particular orthographic strategy of the language. The higher the phonological transparency of the writing system, the less pronounced the effect would be, so the Stroop effect would primarily be expected in ideographic notations. Similar to English, experiments with ASL have also yielded conflicting results: while one study has found it (Vaid & Corina, 1989), no effect was reported in a later study (Bull et al., 2006).

Given the conflicting results of linguistic automatic magnitude processing research, including Number Stroop studies, it has been suggested that the format may fundamentally affect numerical processing (Cohen Kadosh, Cohen Kadosh, Kaas, Henik, & Goebel, 2007; Cohen Kadosh & Walsh, 2009), as opposed to the commonly accepted proposal that there is an abstract, format-independent processing of number (Dehaene et al., 1998). Indeed, numerical notation systems (such as Arabic digits) and number signs/words may represent the same magnitudes, but their use is often governed by different constraints (Chrisomalis, 2019, 2020). Their use in different contexts can also influence processing and retrieval efficiency: for example, doing math problems with written numerals poses more difficulties compared with doing them with digits (Campbell & Alberts, 2009; Campbell & Epp, 2004; Campbell & Fugelsang, 2001), but the skill improves with practice (Metcalfe & Campbell, 2007).

The current study

The conflicting results of two previous ASL Number Stroop effect studies may be related to methodological differences: the experiments had different modes of presentation and stimuli. Vaid and Corina (1989) who found Number Stroop Effect in ASL, presented stimuli sequentially and only used the non-iconic number signs SIX – NINE that are only

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

transparent to those who know the language (the number system of ASL is illustrated in Figure 2).

Bull et al. (2006), on the other hand, presented stimuli simultaneously and used only the number signs ONE – FIVE, which make use of number-to-number iconicity (Taub, 2001) and therefore were transparent and intelligible to the hearing controls as well, but neither group showed evidence of a Number Stroop effect.

Besides the difference between congruent and incongruent trials, Vaid and Corina also analyzed visual field asymmetries (right vs left). They have found a greater Stroop effect in the right visual field for ASL signs and English number words, while the left visual Stroop interference was higher for the digits, which they interpreted as an argument in favor of format-specific number representation. Bull and colleagues approached the possible impact of right/left spatial positioning of the stimuli by analyzing spatial-numerical association of response codes, or the SNARC effect. It is an association of the right side with larger magnitudes and of the left side with smaller ones that is attested in cultures where reading and writing proceeds from left to right (Dehaene et al., 1993), but reversed SNARC effect has been found in cultures reading from right to left (Shaki & Fischer, 2008; Shaki et al., 2009). The presence of the SNARC effect is often interpreted as evidence that numbers (represented by digits or lexemes) may be mapped onto mental number line. However, Bull and colleagues observed some expected trends toward SNARC in numerical judgments, but the effect was not significant in any condition (number or size) or notation (ASL signs or digits).

Additionally, Vaid and Corina discussed language acquisition setting of their participants who learned ASL either early as a first language (deaf people from deaf families, hearing people from deaf families) or later in life as a second language (hearing people from hearing families). Bull et al., on the other hand, did not report the language acquisition background of their participants. Thus, it is possible that the two experiments have detected two different processes in two different groups of people.

Importantly, neither of these two Number Stroop Effect studies reported in their analysis the use of control stimuli, that is, neutral pairs of numbers (such as 3 5 for the number condition, or 3 3 for the size condition), which makes it difficult to evaluate facilitation effects (as opposed to interference effects). To assure that experimental stimuli fully represent the numeral system of ASL, with both iconic (transparent to those who do not know ASL) and non-iconic (nontransparent to those who do not know ASL) number signs, we included all numbers from TWO to NINE, with the number ONE excluded following the original experiment by Henik and Tzelgov (1982) due to its frequency.

To control for age and setting of language acquisition, we compared three groups of participants: first language learners of ASL who acquired language from birth, late first language learners of ASL who first acquired language after the age of 9, and hearing adults acquiring ASL as a second language in a college setting. Doing so allowed us to untangle the effects of age of exposure versus language deprivation: second language

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

and late language learners both began learning ASL late in life, but their prior language experience differed dramatically.

We conducted Number Stroop experiment with two tasks (Arabic digits and ASL number signs) to investigate three questions: whether ASL number signs elicit the Number Stroop Effect, whether this effect is influenced by age of acquisition and/or years of exposure, and whether the effect of age of acquisition is similar for both number formats. In addition, while not the focus of the study, we performed exploratory analyses of possible stimuli effects: the iconicity and frequency of the lexical numerals, and the spatial-numerical congruity of stimuli (the SNARC effect).

Number signs ONE – FIVE are more iconic and more transparent than the signs SIX – NINE. Although it has been shown that iconicity does not facilitate lexical sign processing in native deaf signers (Bosworth & Emmorey, 2010), it may help beginners, or inhibit the processing of experienced second language learners (Baus et al., 2013). Moreover, the cross-linguistic frequency of the first five numbers exceeds the frequency of the subsequent ones (Dehaene & Mehler, 1992), and the frequency of lexemes may influence their recognition and processing (Brysbaert et al., 2017). Spatial positioning of the larger number in a stimulus may also have an effect: the participants of our experiments belong to cultures that write numbers from left to right, where the SNARC effect is expected. Considering the results by Vaid and Corina (1989), it is possible that the SNARC effect can also be format-specific, similarly to the Number Stroop Effect. It is possible, however, that spatial-numerical associations are not strongly activated in a Stroop paradigm, similarly to the results by Bull et al. (2006).

Table 1 lists the possible outcomes of the Arabic Digit task and their potential explanations, and Table 2 lists possible outcomes of the ASL task.

Methods

Participants

Twenty-nine adult users of American Sign Language were recruited. Eleven were hearing second language learners of American Sign Language (all women, mean age (SD): 21.5 (1.08), mean AoA (SD) 15.2 (4.8), mean duration of exposure (SD) 6.1 (5.15)), who acquired ASL in an educational setting (college, university, or high school).

Ten participants were late first language learners of ASL (6 women, 4 men, mean age (SD): 33.1 (12.9), mean AoA (SD): 19.6 (6.12)), mean duration of exposure (SD) 13.4 (14.45). These individuals were born deaf and did not have accessible language input during childhood. Due to various circumstances, these individuals did not have access to natural sign language or spoken language but were not socially deprived. Currently, they are using ASL daily. Two more participants in this group were excluded from the analysis: one did not satisfy the background inclusion criteria (they were exposed to another sign language prior to ASL[1]), and one demonstrated unusually slow reaction times, which suggested that the participant did not perform the task automatically.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Eight participants were deaf early signers of ASL (5 women, 3 men, mean age (SD): 39.7 (13.29)); 7 were exposed to ASL from birth, learning it from their deaf parents, and one participant from a hearing family was exposed to ASL from 1 month of age through an early intervention program. Participants who were not UCSD students received financial compensation for their time, while the students participated in the experiment for class credit (the experimenters were not involved in teaching any of the classes that the extra credit was used for). Participants signed the Informed Consent that was approved by the UCSD Institution Review Board.

Materials

Structure

Each participant performed a computer-based task in two conditions: size comparison (the relevant dimension was physical size) and number comparison (the relevant dimension was number). There were two blocks in each condition: Arabic digits followed by ASL number signs. The order of the size and number conditions was counterbalanced across participants. Within both conditions, stimuli were fully randomized. The experiment was created and performed using the OpenSesame experiment builder (Mathôt et al., 2012).

Each block contained 12 congruent, 12 neutral, and 12 incongruent stimuli, repeated three times with 108 trials per block (ASL or digits) for a total of 216 trials per condition. The structure of each trial was as follows: a white fixation dot appeared in the middle of the black screen for 450 ms, followed by the stimulus (digit/sign array). The stimulus remained on the screen until the participant pressed the key (right or left). Before each block, the participant received instructions in ASL from the experimenter along with explanations from an individual familiar to the participants if needed (in case of late first language learners, a native signer of ASL) and in written English on the screen. In the Arabic digit task, for the number condition, the instructions stated "In this condition, you need to choose a digit that is numerically bigger. To choose the variant on the left, press Z. To choose the variant on the right, press M"; for size "In this condition, you need to choose a digit that is physically bigger. To choose the variant on the left, press Z. To choose the variant on the right, press M." In the ASL condition, "digit" was replaced by "handshape."

Instructions were followed by three examples (congruent, neutral, incongruent): participants saw each example stimulus for 700 ms, after which the correct answer was indicated with green arrows. After the example trials, the participants performed six practice trials followed by feedback, and then began the experiment. Based on the pilot results, to avoid boredom that may lead to inadequate effort on cognitive tasks completed exclusively for credit (DeRight & Jorgensen, 2015) and increase motivation, each block was followed by feedback as well: the percentage correct and mean reaction time (RT). The participants were encouraged to respond as fast as possible.

[1]The background criteria for late first language learners included being born deaf, not being exposed to sign/spoken language prior to the age of 10, and not being socially deprived.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Stimuli

The Arabic digits 2–9 and the ASL signs TWO – NINE were used as stimuli. Each digit/number sign was paired with itself for a neutral comparison in physical size, or with a different number that was always numerically smaller or bigger by two (for example, 5 was paired with 3 or 7). The signs/digits differed in physical size and numerical magnitude. The bigger item size was 3.2", the smaller item size was 2.9". Digit stimuli were created using standard font Calibri (Body). The white stimuli were presented on a black background based on the suggestions from the pilot subjects. The ASL handshape illustrations were created from photographs of a native signer signing numbers. Examples of the stimuli for each block are shown in Figures 3 and 4. To avoid right/left hand biases, each digit/number sign appeared on each side of the screen an equal number of times.

Results

Age

Given the small size and heterogeneity of the groups of participants in terms of age, we first explored whether age influenced the overall reaction time (RT) independently of the Stroop interference and language acquisition circumstances, since several studies have suggested that the Color Stroop Effect changes with chronological age, namely participants who are older generally respond more slowly (Bugg et al., 2007; West & Baylis, 1998), although other studies contest this effect (Verhaeghen & De Meersman, 1998).

For each numerical format (digits and ASL), we built linear regression models (using *lm* function in *R* (R CORE TEAM, 2016)) with mean reaction time for each participant as a dependent variable and age of participant as a predictor variable. The effect of age was not significant neither for Arabic digits ($\beta = 8.495$, $CI = -0.66$–17.66, $SE = 4.464$, $t(29) = 1.9$, $p = .07$) nor for ASL number signs ($\beta = 4.61$, $CI = -5.97$–15.19, $SE = 5.517$, $t(29) = 0.894$, $p = .379$).

Number Stroop Effect: data processing

The experimental within-subject factors were size comparison (physical vs. semantic), notation, i.e., format (Arabic Digits vs ASL number signs), congruity (congruent, incongruent, neutral), condition (number and size). The between-subject factors were block order (size or number first) and Age of acquisition (AoA). All the variables were categorical (for this analysis, AoA included three groups – early first language learners, late first language learners, second language learners).

Data analyses for response time were conducted for correct response trials only. The outliers for each subject were removed using an interquartile rule 1.5 x (IQR). Some previous Number Stroop Effect studies have used a cutoff method and included only the trials with reaction times under a specified threshold (for example, 150–2000 msec) in the analysis (Cohen Kadosh et al., 2011; Szucs & Soltész, 2007). However, we did not use a cutoff method here because, in relatively small sample sizes with large variation, as in the present study, a general threshold may affect the power and introduce asymmetric biases (Whelan, 2008). There was a high degree of individual variation within our sample, especially for the

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

late first language learners. In the following sections, the results for each task are presented separately, first for the Arabic Digit task, then for the ASL number sign task.

Task 1: Arabic digits

Accuracy

Overall accuracy was high for all groups, with the late first language learners showing somewhat lower accuracy, that was still above chance (Early language learners: 0.96 (SD 0.03), Second language learners: 0.95 (SD 0.02), Late first language learners: 0.88 (SD 0.11)). To further explore this difference between groups, we used a linear mixed-effect regression model that included accuracy as a dependent variable and condition (size vs number), congruity (congruent, neutral, and incongruent), and Age of Acquisition and all their interactions as predictor variables, and a random intercept for participants. No main effect was significant, but the interaction between congruity and age of acquisition was significant: late first language learners were significantly less accurate on incongruent trials ($\beta = -0.20$, $CI = -0.34 - -0.05$, $SE = 0.07$, $t(174) = -2.4$, $p = .007$). Their group performance on incongruent trials was still above chance (mean LL1 accuracy in number condition on incongruent trials 0.72, SD = 0.023, binomial probability test: $CI = 0.627$–0.804, $p < .001$; mean accuracy in size condition 0.89, SD = 0.23, binomial probability test: $CI = 0.802$–0.934, $p < 0.001$). Two participants that completed size condition first, performed with high accuracy in size condition but below chance in number condition, and one participant who completed the number condition first, showed high accuracy in number condition and below chance accuracy in size. Mean reaction time with accuracy scores for each group for each congruity level in the Number Stroop task in Arabic digits are presented in Table 3.

Reaction time

The deaf early first language learners and hearing ASL L2 learners showed comparable performance in terms of speed. By contrast, the mean RT for the late first language learners was slower (Table 3).

To estimate congruity effects in both conditions (size/number), we performed a mixed-effects regression model in R, using the package *lme4* (Bates et al., 2016). The predictor variables were the within-participants factors of congruity (congruent, neutral, incongruent) and condition (size, number), and the between-participants factor was Age of Acquisition (early first language learners, late first language learners, hearing second language learners). The interactions of congruity and condition with AoA were included in the model. We included random intercepts for block order, number or size first, and participants (nested) and stimuli (every stimulus was seen by each participant three times). The model was tested for multicollinearity (for all effects VIF < 3.5). Confidence intervals were verified through the *confint ()* function with a bootstrapping resampling technique, based on 1000 bootstrapping replicates. All the significant effects were confirmed, so we report the CI obtained through bootstrapping.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

We first fit the model that included Condition, Congruity, and AoA with no interactions as predictors, followed by a model that included interactions of AoA with condition and AoA with congruity. We compared these two models based on the results of previous studies. In adults, congruity effects with Arabic digits have been shown reliably across populations. At the same time, in children, the emergence and nature of the congruity effect changes with the amount of number exposure (Girelli et al., 2000; Heine et al., 2010; Rubinsten et al., 2002). Difference in RT between size and number conditions may also change with the amount of exposure, and therefore age and setting of language exposure might influence both the congruity and condition effects differently across groups. Since the Akaike Information Criterion (estimator of out-of-sample prediction error) was lower for the second model including interactions (76805 and 76792), the analysis was performed using this model. The graph representing reaction time for the Arabic digit task is shown in Figure 5. The results of the model are presented in Table 4.

The main effect of congruity was significant for both facilitation (congruent being faster than neutral, $\beta = -31.98$, $CI = -58.7 - -7.95$, $SE = 13.04$, $t(5728) = -2.45$, $p = .014$) and interference (incongruent being slower than neutral, $\beta = 49.94$, CI = 23.73–75.34, SE = 13.15, t (5728) $= 3.79$, $p < .001$)). The main effect of comparison condition was also significant, with size judgments being faster than number judgments ($\beta = -158.63$, $CI = -176.34 - -139.53$, $SE = 9.28$, $t(5728) = -17.094$, $p < .001$).

Differences in the Stroop effect between groups

While the reaction time of the deaf early first language learners (L1) and hearing second language ASL learners (L2) groups did not significantly differ, the late first language learner group (LL1) demonstrated significantly slower reaction time ($\beta = 306.88$, $CI = 126.42$–469.98, $SE = 85.00$, $t(5728) = 3.610$, $p < .001$). Moreover, the interaction between condition and age of acquisition (AoA) was also significant: the mean difference in RT between size and number judgments for the late first language learner group was significantly larger than it was for the early first language learners ($\beta = -55.63$, $CI = -81.72$–30.19, $SE = 12.939$, $t(5728) = -4.299$, $p < .001$).

Size of interference effect

Number Stroop Effect in terms of interference was found in all groups. To determine if the size of such interference varies as a function of age of acquisition, we performed an additional linear mixed effect regression model with the mean difference between RT for neutral and incongruent stimuli for each participant as the dependent variable. The predictor variables were age of acquisition (L1, L2, LL1), condition (size and number) and their interactions. The main effect of age of acquisition was significant with late first language learners showing a larger difference between neutral and incongruent trials than the early first language learners ($\beta = -238.58$, $CI = -421.66 - -55.49$, $SE = 91.238$, $t(158) = -2.615$, $p = .012$). No other effect was significant.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Random effects

Stimuli: SNARC effect and perceptual similarity

Since the effect of stimuli was significant (CI obtained by bootstrapping 11.119–26.631), we performed additional analyses to evaluate if the difference in RT was caused by the structure of the stimuli that elicited the Spatial–Numerical Association of Response Codes, or the SNARC effect (Dehaene et al., 1993). In cultures that write numbers from left to right, people tend to react faster to larger numbers that require rightward response, and to smaller numbers that require leftward response (Fias, 2001; G. Wood et al., 2008). Several studies suggest that the SNARC effect depends both on left to right (or right to left) reading habits (Shaki & Fischer, 2008; Shaki et al., 2009) and immediate spatial experiences (Fischer et al., 2009). While the SNARC effect is usually assessed through number parity judgments without size incongruities involved, there was a possibility that it can influence the processing times for particular stimuli.

Traditionally, stimuli for SNARC effect only have one dimension of congruity: 3 5 can be an example of the congruent stimuli where the right number is smaller in both dimensions. However, in a Stroop-like tasks the stimuli varied in congruity in two dimensions. Thus, a stimulus where 7 is physically smaller than 5 is incongruent numerically, but congruent spatially (right number being bigger in size).

For the purpose of the subsequent analysis, we defined the stimuli as numerically SNARC-congruent if the right number was numerically bigger than the left one; size SNARC-congruent if the right number was physically larger , and overall SNARC-congruent if numerical and size information aligned in terms of the SNARC effect (for example, when the right number is bigger in both size and number, or the right number is smaller in both dimensions). The stimuli that only had one dimension of comparison (neutral) were excluded from the analysis.

Using *lm* function in *R* (R CORE TEAM, 2016), we built a linear regression model with reaction time for each stimulus as a dependent variable. Numerical SNARC congruity, size SNARC congruity, group (early first language learners, late first language learners, or hearing second language learners of ASL), and their interactions were used as predictor variables. The main effect of size SNARC congruity was significant with size SNARC-incongruent stimuli being processed more slowly ($\beta = 173.46$, $CI = 30.87–316.04$, $SE = 72.08$, $t(144) = 2.40$, $p = .017$). The main effect of group was also significant: the late learners of ASL were significantly slower than other groups ($\beta = 308.69$, $CI = 154.69–462.70$, $SE = 77.86$, $t(144) = 3.96$, $p < .001$). The interaction between two types of SNARC effect (or, as we define it, overall SNARC effect) was significant: stimuli where size and number congruity/incongruity aligned were processed faster than stimuli that are incongruent/incongruent in one dimension ($\beta = -241.99$, $CI = -442.58 – -41.41$, $SE = 101.40$, $t(144) = -2.38$, $p = .018$). The overall SNARC effect is illustrated by Figure 6.

Additionally, the stimuli including the digits 6 and 8 as SNARC-congruent were processed 40 ms slower than the baseline. It has been suggested in previous literature that processing speed for larger and smaller numbers might differ due to the magnitude or frequency effects

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

(Girelli et al., 2000; Tzelgov et al., 1992). Using the linear regression model, we analyzed whether mean reaction times for the stimulus depended on magnitude (small (1–5), large (6–9) or mixed (stimuli containing both)), but the effect of magnitude was not significant: stimuli with neither small ($\beta = -59.67$, $CI = -157.30–37.96$, $SE = 49.38$, $t(144) = -1.208$, $p = .229$) nor large magnitudes ($\beta = -21.07$, $CI = -118.69–76.56$, $SE = 49.38$, $t(144) = -0.427$, $p = .67$) were processed differently from the mixed magnitudes stimuli. However, it has been previously shown that perceptual similarity between digits can significantly influence the speed of their discrimination, and eight differs from six with only one line compositional element (Cohen, 2009), and this might explain the difficulty of distinguishing 6 and 8 specifically.

Individual differences: delayed first language acquisition and language experience

The nested random effect of order/participant was significant (CI obtained by bootstrapping 134.43–232.77). Since participants in the late first language acquisition group varied greatly in their age of acquisition and years of exposure, we analyzed the potential impact of these factors on the reaction times. We built a linear regression model (using *lm* function in *R* (R CORE TEAM, 2016)) with mean reaction time for each participant as a dependent variable. The predictor variables were the exact age of first language acquisition (AoA), number of years of exposure (YoE) and their interaction. For this analysis, AoA and YoE were continuous variables. Only the main effect of years of exposure was significant ($\beta = 64.96$, $CI = 5.70–124.22$, $SE = 24.21$, $t(10) = 268$, $p = .036$): the more years of experience the late learners had, the slower they were. This somewhat surprising effect is addressed in more detail in the Discussion.

Summary of the Task 1 results

Overall, the results showed the expected size congruity Number Stroop effect: both interference and facilitation were shown for the number judgment task, and interference only for size judgment, in line with previous studies (Girelli et al., 2000; Henik & Tzelgov, 1982). The Condition effect (size, number) was significant as well. However, the effects differed depending upon the group. Deaf and hearing participants who learned a first language early in life performed identically. In contrast to the deaf early signers, deaf participants who experienced highly delayed exposure to language showed slower RT, larger RT differences between number and size judgments, and greater interference effect (i.e. RT difference between congruent and neutral trials and reduced accuracy on incongruent trials). With more years of language experience, late first language learners did not become faster, but instead demonstrated the tendency toward slower reaction times. Exact age of first language acquisition did not correlate with the processing speed.

Additionally, all groups demonstrated a size-SNARC congruity effect and the overall SNARC effect (for stimuli where numerical and size information aligned in terms of congruity). Unlike with the Stroop, there were no differences between groups in the SNARC effects. Other characteristics of stimuli (frequency and magnitude size) did not significantly affect reaction times.

Author Manuscript

Together, the results of Task 1 suggest that early language deprivation affected automatic magnitude processing, but magnitude processing was still achieved despite impoverished early input (i.e., only digits, but no language).

Task 2: ASL number signs

Accuracy

Overall accuracy was high for all groups, with the late first language learners showing somewhat lower accuracy (ASL: L1 0.96, L2 0.96, LL1 0.92). We fit a linear mixed-effect regression model that included accuracy as a dependent variable and condition (size vs number), congruity (congruent, neutral, and incongruent), and age of acquisition with all their interactions as predictor variables, and a random intercept for participant. No main effect was significant, but the interaction between congruity and age of acquisition was significant. Late first language learners were significantly less accurate on incongruent trials ($\beta = -0.17$, $CI = -0.30 - -0.03$, $SE = 0.068$, $t(174) = -2.4$, $p = .015$), but as a group still above chance (mean LL1 accuracy in number condition on incongruent trials 0.78, SD = 0.39, binomial probability test: $CI = 0.687–0.852$, $p < .001$; mean accuracy in size condition 0.92, SD = 0.15, binomial probability test: $CI = 0.847–0.961$, $p < .001$). Two participants that completed size condition first, performed with high accuracy in size condition but below chance in number condition, and one participant who completed the number condition first, showed high accuracy in number condition and below chance accuracy in size. These participants exhibited the same pattern in Arabic digit task. Mean reaction times and accuracy scores for the ASL number signs are shown in Table 5.

Reaction time

RT differed greatly between groups and conditions, although size judgments were made at comparable speed by deaf early first language learners and hearing second language learners.

Following the same rationale described above for Task 1, we first fit the model that included Condition, Congruity, and AoA with no interactions as predictors, followed by a model that included interactions of AoA with condition and AoA with congruity (both models included random intercepts for stimuli and participant/block order (nested). Since Akaike Information Criterion for the model with interactions was smaller (75,000 and 74,729), it was used for the subsequent analysis. Multicollinearity was checked through VIF (all VIF < 2.5). The ASL RT is shown in Figure 7, and the full results of the model are shown in Table 6.

In contrast to the Arabic Digit task, the main effect of congruity was not significant when magnitudes were represented by ASL signs. However, there was an interaction effect of congruity with age of acquisition: in the late first language group, the neutral stimuli were processed significantly more slowly than the congruent stimuli ($\beta = -40.26$, $CI = -78.68 - -2.58$, $SE = 19.94$, $t(5379) = -2.019$, $p = .04$). Differences in congruity were not significant for any other group.

However, the main effect of condition (size, number) was significant: size judgments were faster than number ($\beta = -385.584$, $CI = -410.99 - -361.42$, $SE = 12.448$, $t(5379) =$

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

−30.976, p = <.001). The main effect of age of acquisition group was significant as well: both the hearing second language learners and deaf late first language learners significantly differed from the early deaf first language learner group (L2: β = 221.80, CI = 33.48–425.65, SE = 94.06, t(5379) = 2.358, p = .018, LL1: β = 302.93, CI = 118.53–494.29, SE = 94.00, t(5379) = 3.223, p = .001). In the size condition, mean reaction times of the hearing second language learner group were very close to those of the early first language group, but in the number condition the hearing second language learners performed as slowly as the late learners.

The significant interaction between condition and age of acquisition indicated that all three groups showed contrasting RT patterns as a function of size and number. The largest difference in performance between the number and size conditions was shown by the hearing second language learners (β = −190.91, CI = −226.04 – −160.40, SE = 16.732, t(5379) = −11.410, p = <.001). By contrast, the smallest difference in performance between the number and size conditions was shown by the late first language learners, due to their slowed performance in the size condition (β = 84.03, CI = 49.44–115.34, SE = 16.586, t(5379) = 5.067, p = <.001).

Random effects

Stimuli: ASL SNARC effect and iconicity

Since the random effect of stimuli was significant (CI from bootstrapping 40.454–67.663), we performed an additional analysis identical to the one described in Task 1 to detect a possible SNARC effect and its interaction with language acquisition group. However, none of the SNARC effects (number, size or overall) was significant (size SNARC: β = 92.51, CI = −165.76–350.77, SE = 130.56, t(144) = 0.709, p = .047, number SNARC β = 142.28, CI = −120.09–404.66, SE = 132.64, t(144) = 1.073, p = .285), and the only significant result was that late learners demonstrated slower reaction times (β = 355.89, CI = 76.93–634.86, SE = 141.03, t(144) = 2.524, p = .013). None of the interactions were significant. Another potential source of variation can be the transparency of the stimulus, or whether it abides to number-to-number iconicity.

The combinations of number signs in our stimulus set can be divided in three groups: only iconic transparent numerals (THREE FIVE, TWO FOUR, FIVE THREE, FOUR TWO), a mix of transparent and nontransparent (FOUR SIX, FIVE SEVEN, SIX FOUR, SEVEN FIVE), and nontransparent (SIX EIGHT, SEVEN NINE, EIGHT SIX, NINE SEVEN). To evaluate the effect of this transparency, we used the *anova* function of R to compare two linear regression models. One included mean reaction time for the particular stimulus as a dependent variable and Stroop congruity, age of acquisition group, and condition as predictor variables, the other model also included transparency of the stimulus (transparent, nontransparent, or mixed); the model that included transparency had better R^2/R^2 adjusted. Table 7 presents the results of the models. Alongside the main effects of group and condition, the main effect of number transparency was significant: the stimuli with transparent (iconic) number signs TWO to FIVE were processed faster than the mixed stimuli that included a combination of transparent and nontransparent number signs, but the stimuli with nontransparent signs SIX to NINE did not differ from the mixed stimuli.

However, there is a possibility that the effect was produced not by transparency, but by higher crosslinguistic frequency of the first five numbers (Dehaene & Mehler, 1992): the nontransparent number signs in ASL all designate higher magnitudes that are less frequent.

Individual differences: delayed first language acquisition and language experience

Since the random effect of participant was significant (CI obtained by bootstrapping 149.482–255.979), we performed additional analyses to compare the influence of age of ASL acquisition and Years of Exposure on number sign processing in deaf late learners and hearing second language learners of ASL. The linear regression model included reaction time as a dependent variable and exact age of acquisition, exact years of exposure (both were continuous variables), and their interactions. The only effect that was significant was years of exposure ($\beta = 52.57$, $CI = 4.34–100.80$, $SE = 22.860$, $t(21) = 2.300$, $p = .034$); participants demonstrated high variation in reaction time patterns, but in both groups, there were several individuals with longer exposure to ASL who performed slower than people with comparable or less exposure. This effect will be addressed in discussion.

Summary of the Task 2 results

We did not find the typical Number Stroop Effect in the ASL condition. Age and setting of ASL acquisition also impacted performance: while in the size condition second language learners performed no differently from early signers of ASL, in the number condition they were significantly slower. Late learners, on the other hand, were slower in both conditions and, similarly to the Arabic digit task, demonstrated decreased accuracy on incongruent trials. Hence, unlike in the Arabic Digit task, the largest difference between size and number judgments was demonstrated by hearing second language learners of ASL, who still performed with high accuracy.

The SNARC effect was not attested in the ASL condition as well, but another effect of stimuli was significant: stimuli with frequent and transparent number signs TWO to FIVE were processed faster than stimuli with less frequent nontransparent number signs and mixes of transparent and nontransparent ones.

Some of the deaf late learners and hearing second language learners of ASL demonstrated a tendency toward slower reaction times despite their longer experience with the language. Exact age of first language acquisition did not correlate with the processing speed.

Discussion

In the current study, we conducted a Number Stroop experiment with two tasks (Arabic digits and ASL number signs), with three groups of participants (deaf early learners, deaf late learners, and hearing second language learners of ASL) to investigate three questions: whether automatic magnitude estimation is influenced by age of acquisition and/or years of exposure, whether the Number Stroop Effect is found in ASL number signs as well, and whether the effect of age of acquisition is similar for both number formats.

Revisiting the possible outcomes in Tables 1 and 2, in Task 1, the results showed the Number Stroop Effect with Arabic digits was present in all groups, but there were

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

differences in late first language learners, suggesting that age of acquisition might affect automatic magnitude processing, but it can still be achieved despite incomplete early input.

The Results of Task 2 suggested that, since the Number Stroop Effect in ASL was not found in any group, ASL number signs activate magnitudes in a different way from Arabic digits, supporting the format-specific activation hypothesis. At the same time, both late and second language learners differed from the early first language learners, suggesting that both years of exposure and age of acquisition influence automatic magnitude processing with number signs.

The results of the two tasks are discussed separately, followed by the general discussion.

Magnitude estimation and age of acquisition: Arabic digits

The results showed the expected Number Stroop effect (incongruent stimuli were processed more slowly than neutral, and congruent were faster than neutral in number condition) and condition effect (the size comparison was faster than the number comparison) in all groups, but age of acquisition influenced the results. Deaf and hearing participants who learned language early in life performed identically, but late first language learners showed slower RT, a larger time difference between number and size judgments, a larger difference between incongruent and neutral trials, and lower accuracy on incongruent trials.

The large difference in speed between size and number judgments was previously attested in first-graders who successfully completed various number tasks, including counting and matching Arabic digit to the correct numerosity (Girelli et al., 2000). However, these young children also did not show the canonical Number Stroop Effect. There was no interference in size condition, and incongruity in number condition affected response accuracy (it was lower) but not reaction time (it was similar to RT for neutral stimuli). These first-graders were then compared with third- and fifth-graders who demonstrated a more adult-like pattern of Stroop Effect (Girelli et al., 2000; Rubinsten et al., 2002). The authors interpreted the result as evidence for developmental changes in integration of size and number information and gradual automatization of magnitude processing which comes with experience.

In contrast, in the present study late first language learners demonstrated a robust Number Stroop effect in the number task and experienced even greater interference (i.e. difference between neutral and incongruent trials) than early signers of ASL and hearing second language learners. Late first language learners also demonstrated lower accuracy on incongruent items. This result suggests that the difference between size and number conditions in late first language learners and in children requires different explanations. First graders may have not yet fully achieved automatic magnitude processing and integration of size information with numerals, as this integration develops with experience. The fact that late first language learners demonstrated a robust Stroop effect in number judgments suggests that both dimensions are salient for them. It is possible that late learners might instead experience greater difficulties suppressing irrelevant information. This effect is exacerbated by the task switch: three participants were able to complete their first condition

Author Manuscript

(size or number) with high accuracy, but when the task changed to the opposite one, their accuracy dropped to the below chance level, despite the successfully completed practice.

The difference between young children and late first language learners is underscored by the fact that with more years of language experience, late first language learners did not become faster, but demonstrated a tendency toward slower reaction times. This result suggests that while more exposure leads to automaticity of magnitude processing (and a stronger Stroop effect), delayed first language acquisition may affect the inhibition of irrelevant information and thus slow down the decision and affect accuracy. However, taking into account the small sample and the variety of life experiences of the participants, this result needs to be interpreted with caution. Exact age of first language acquisition did not correlate with processing speed, suggesting that the effect of language deprivation is not gradual after early childhood, but abrupt, in line with previous research, showing the absence of correlation between exact age of acquisition and performance on linguistic and cognitive task battery (Mayberry, Hatrak, Ilkbasaran, Cheng, & Hall, in prep).

Additionally, all groups demonstrated an overall SNARC congruity effect and a size SNARC effect, but not a numerical SNARC. Note that canonically, the SNARC effect is studied with number comparison or parity judgment tests, but not Stroop-like paradigms, and therefore this result might be a byproduct of the particular methodology. For instance, one previous study did not find significant SNARC effects in a different Stroop paradigm in both hearing and deaf participants (Bull et al., 2006). On the other hand, another study did find the SNARC effect in deaf individuals in a number comparison task, but with slower reaction times (Bull et al., 2005). Unlike the Stroop effect, there were no differences between groups in the SNARC effects in our study, with both late and early deaf signers of ASL experiencing the same effect as the hearing participants.

Together the results of Task 1 suggest that early language deprivation may affect automatic magnitude processing, but that it still can be achieved despite incomplete early input. When language is acquired on a typical timeline, deaf participants score identically to the hearing participants, challenging the results of the studies that link a slowdown in number processing to hearing loss itself.

Number Stroop Test in ASL: no Stroop effect

The typical Number Stroop Effect was not attested in the ASL task. Predictably, age and setting of ASL acquisition impacted performance: while in the size condition the second language learners performed with no differences compared with the early signers of ASL, in the number condition they were significantly slower, although their accuracy was still high. This may indicate more careful decision process related to a lack of proficiency. Late learners, on the other hand, were slower in both conditions and were less accurate on incongruent trials in numerical comparison. Similarly to the Arabic digit task, three participants were impacted by the task switch and were able to complete their first condition (size or number) very accurately but dropped to chance level, once the dimension of comparison changed. These participants understood the task and completed practice successfully, but during the test it was hard for them to overcome interference from irrelevant dimension enhanced by the first task.

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

Additionally, late learners of ASL demonstrated the slowest reaction times for neutral stimuli in the size condition, which is an unusual pattern that has not been described in previous studies. Previous studies (using digits) with participants with developmental dyscalculia have reported abnormal patterns, but these effects were related to the facilitation effect patterns (Ashkenazi et al., 2008), without a slow down on neutral stimuli. The comparison in question involved pictures of the same number handshapes (for example, two FIVE handshapes) that only differed in size; the numerical difference was not present at all. We hypothesize that late language learners might experience difficulties because, of all comparisons on the test, this one is the most unusual. While people do in fact see number words and Arabic digits written with various contrasting font sizes in real life (for example, in advertising), this doesn't happen with sign language perception: signers' hands do not change size, and the contrast between photos is perhaps not as salient as with printed digits. Other groups might have adapted to the unusual task easier than the late learners of the present experiment.

SNARC effects were not attested in the ASL condition as well. However, previous studies have identified numerical SNARC effect in German (DGS) and Italian (LIS) sign languages, using parity judgment tasks (Bull et al., 2005; Chinello et al., 2012; Iversen et al., 2006, 2004). We attribute the difference with our results to the experimental paradigm: the Stroop paradigm is less efficient for detection of spatial association of magnitudes. Since there are two interacting dimensions of SNARC congruity (size and number), the canonical numerical only SNARC effect may not be assessed reliably. Indeed, another Stroop paradigm study with ASL number signs did not report a significant SNARC effect either (Bull et al., 2006). Alternatively, the explanation might relate to the structure of the numeral system: LIS and DGS have two-handed numeral systems, and in these languages the compositional structure of two-handed numerals has a sub-base of 5, which influenced parity judgments. In two-handed number signs, the non-dominant hand has the same handshape (FIVE), while handshape on the dominant hand changes, and there is a direction of sign perception than can be compared to the direction of reading. ASL number signs are one-handed.

Another effect of the stimuli was significant: stimuli with frequent and transparent number signs TWO to FIVE were processed faster than stimuli with the less frequent nontransparent number signs and mixes of transparent and nontransparent ones. The difference in RT can be attributed either to iconicity or to the frequency of the first five numbers, since their frequency crosslinguistically exceeds the frequency of the subsequent ones (Dehaene & Mehler, 1992). There are two arguments in favor of the frequency hypothesis. The frequency ratings from the ASL-Lex database (Sehyr et al., 2021) confirm that for ASL, this relationship also holds. Moreover, a similar effect (with faster reaction times for smaller numbers) was found in Italian Sign Language, which has a fully iconic and transparent two-handed numeral system (Chinello et al., 2012). This is another argument in favor of frequency but not iconicity being a facilitating factor. Finally, if iconicity alone was in play, then mixed stimuli would also be processed faster, since all nontransparent number signs refer to larger magnitudes than transparent iconic ones, and there would be no need to even interpret them to answer the question of which is larger, and yet it does not facilitate the decision.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Finally, we examined whether the exact age of ASL acquisition and exact number of years of experience influenced the processing of ASL numbers in late and second language learners. Exact age of ASL acquisition for late learners did not correlate with the processing speed, suggesting the existence of the critical period. Beyond early childhood, the exact age of language acquisition does not have a significant effect, in line with the result previously shown by Mayberry et al. (in prep). Success of second language learning may not depend on age of acquisition as well. We found a significant effect of years of exposure, but, similar to the digit condition, it is the opposite of what one might expect: some of the late learners and second language learners of ASL demonstrated a tendency toward slower reaction times despite their greater experience with the language. An explanation might be related to the life experience of participants: both second language learners who were currently acquiring ASL in a classroom setting and the late learners who were immersed in the Deaf community and were taking ASL or English classes more recently, might have more fresh experience with timed tasks and therefore perform faster than participants that had this experience longer ago. However, the small sample and the variety of life experiences of the participants are serious limitations to this generalization.

Overall, the results of Task 2 show that magnitude activation by ASL number signs and Arabic digits differs. Similar results have been obtained for spoken languages with non-ideographic writing systems, such as Hebrew, Hindi, and Japanese when written with syllabic script (Besner & Coltheart, 1979; Cohen Kadosh et al., 2008; Cohen Kadosh & Walsh, 2009; Takahashi & Green, 1983; Vaid, 1985). The significant difference between the language background groups suggests that both years of exposure and age of acquisition influence automatic magnitude processing with number signs. There was no gradual effect of age of acquisition in late first language learners: if the language was learned post childhood, the outcomes were similar. However, overall high accuracy demonstrated by both second and late ASL learners shows that ASL numbers were successfully acquired by both groups, but late first language learners experience more difficulty suppressing the interference of irrelevant information, similarly to the results of Task 1.

General discussion

Together the results of the two tasks suggest that magnitude information is accessed differently depending on the format (number signs or digits). The results further show that late first language learners can acquire and use both formats. However, their ability to do so is affected by language deprivation in both formats. While some specific patterns of late first language learners' performance appear to be format-specific (a large difference between size and number in the digit task, with the longest reaction times for the neutral stimuli in the size condition in ASL task), this group performed slower in both formats and was more affected by the interference of irrelevant information in terms of accuracy.

It has been shown that late first language learners performed more slowly than early ASL signers in various ASL tasks, but faster than second language learners, or at a comparable speed (Ferjan Ramirez, Leonard et al, 2013; Ferjan Ramirez et al., 2016; Mayberry, Davenport, Roth, & Halgren, 2018), and their performance in non-verbal cognitive tasks is comparable to hearing controls (Mayberry et al., 2018). Therefore, we hypothesized

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

that the slow performance in our experiment was not a general property of the late first language learner group, but may represent the specifics of their magnitude processing. Slower reaction time may be associated with difficulties inhibiting irrelevant information – but it could also be associated with educational deprivation and little experience with timed tasks, although by the time of testing all the late first language learners had already had the educational experience of a classroom setting, taking exams, and playing games where time and reaction are important. The effects of language deprivation and educational deprivation are hard to untangle, since one inevitably creates the other. However, the finding that delayed first language deprivation may be associated with slower response times on magnitude processing tasks may help explain the conflicting results of earlier studies. Effects of language acquisition setting are often not controlled (see Hall and Dills (2020) for a detailed analysis of this issue) and may be relevant for the interpretation of studies that report a slowdown in magnitude tasks in deaf people (for example, Bull et al., 2005; Epstein et al., 1994).

At the same time, in comparison to the detrimental effects of early first language deprivation on language proficiency that have been described in the literature (Boudreault & Mayberry, 2006; Cheng & Mayberry, 2021, 2019; Fromkin et al., 1974; Mayberry et al., 2017; Mayberry et al., 2018; Newport, 1990), the acquisition of basic numbers appears to be more intact: late first language learners perform with overall high accuracy with Arabic digits, and they demonstrate strong evidence of automatic magnitude activation, typical of adults in a numerical culture. With ASL number signs, late first language learners demonstrate even higher accuracy than with digits.

What makes numbers so special? Perhaps, the numerical culture that the participants live in makes number so fundamental that, despite the absence of conventional language input, from an early age the late learners still use quantities, rely on numbers, watch people use number gestures and communicate number information to them. The studies of homesigners in Nicaragua, another example of a highly numerate culture, documented quantity-tracking devices emerging in homesign systems without language models (Coppola et al., 2013), even though these devices function more similarly to indices of items within sets rather than cardinal representations of sets (Spaepen et al., 2013), and conventional signs for large exact numbers may not be developed (Spaepen et al., 2011). While it has been shown that a counting list is needed to form the representation of larger numerosities, the concept of exactness is engrained in the numerical culture in which late first language learners grew up. Besides language, number development also requires the approximate number system to be intact. Finally, our experiment only assessed automatic number representation, and more research is needed to establish how language deprivation affects more complex mathematic operations.

The results of the Number Stroop Test with ASL numerals did not reveal a Number Stroop Effect in any age of acquisition group. These results are in line with the results by Bull et al. (2006), but not those of Vaid and Corina (1989). This may be due to methodological differences. As discussed earlier, Vaid and Corina presented their stimuli sequentially, while our experimental procedure included simultaneous presentations of stimuli, as in Bull et al. (2006). This might indicate that, due to differences in experimental design, these studies detect different automatic processes. The absence of Number Stroop Effect in simultaneously presented linguistic Stroop stimuli is in line with the results of several experiments on spoken languages and supports the hypothesis that mechanisms of automatic magnitude processing may be format-dependent (Cohen Kadosh & Walsh, 2009). According to this hypothesis, the processing of linguistic numerals may be less automatic even if unintentional, because it requires more processing resources. This prevents interference from size information. Neuroimaging research suggests some format-specific differences in processing as well (Cohen Kadosh et al., 2007).

Although automatic magnitude processing by linguistic numbers produces reaction time patterns that differ from the Stroop Effect observed with Arabic digits, the decreased speed of magnitude processing in late first language learners suggests a link between the two formats of number representation. However, in line with research conducted in Nicaragua with deaf and hearing adults of various backgrounds (Flaherty & Senghas, 2011), numbers can be successfully acquired later in life. Despite being more prone to interference and showing an unusual speed and pattern of reaction time, the late first language group demonstrated automatic magnitude activation, which is needed for skilled calculation.

Together, the data from both formats (digits and linguistic numerals) suggest that early first language exposure matters for number acquisition, and when language is acquired early in life, its modality does not have an effect on number representation: deaf early signers are as fast and accurate as hearing controls. This result once again underscores the importance of early access to natural sign languages for all deaf children. Our results also call for adequate control for language background in studies of deaf education: when ignored, the effect of language deprivation can be confounded with other factors.

Acknowledgments

This work was supported by the National Institute on Deafness and other Communication Disorders (NIDCD) (R01DC012797, PI Rachel I. Mayberry), National Science Foundation (NSF) (Linguistics Program - Doctoral Dissertation Research Improvement Award, 19141456, PI Rachel I. Mayberry, co-PI Nina Semushina), Annette Merle-Smith Fellowship for graduate students in anthropogeny, and Malcolm R. Stacey Memorial Fellowship (both to Nina Semushina). The content of the research is solely the responsibility of the authors and does not necessarily represent the views of NIH. We are grateful to Marla Hatrak, Peggy Lott and Kay Vincent who assisted with subject recruitment, to Qi Cheng who assisted with data collection, to Tory Sampson, Deniz İlkbaşaran, Agnes Villwock, Beatrijs Wille, and Eva Wittenberg for the feedback on the stimuli and experiment design, and to Leon Bergen and Semion Leyn for the feedback on earlier versions of statistical analysis. We also wanted to thank the editor and three anonymous reviewers who helped us to greatly improve the manuscript, and all deaf and hearing participants who made this research possible.

Funding

This work was supported by the National Science Foundation, Malcolm R. Stacey Memorial Scholarship; Annette Merle- Smith Graduate Scholarship Award; the National Institute on Deafness and Other Communication Disorders.

References

Algom D, Dekel A, & Pansky A (1996). The perception of number from the separability of the stimulus: The Stroop effect revisited. Memory and Cognition, 24(5), 557–572. 10.3758/BF03201083 [PubMed: 8870527]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Anderson A, Anderson J, & Shapiro J (2005). Supporting multiple literacies: Parents' and children's mathematical talk within storybook reading. Mathematics Education Research Journal, 16(3), 5–26. 10.1007/BF03217399

Arfé B, Lucangeli D, Genovese E, Monzani D, Gubernale M, Trevisi P, & Santarelli R (2011). Analogic and symbolic comparison of numerosity in preschool children with cochlear implants. Deafness & Education International, 13(1), 34–45. 10.1179/1557069X11Y.0000000002

Ashkenazi S, Henik A, Ifergane G, & Shelef I (2008). Basic numerical processing in left intraparietal sulcus (IPS) acalculia. Cortex, 44(4), 439–448. 10.1016/j.cortex.2007.08.008 [PubMed: 18387576]

Ashkenazi S, Rubinsten O, & Henik A (2009). Attention, automaticity, and developmental dyscalculia. Neuropsychology, 23(4), 535–540. 10.1037/a0015347 [PubMed: 19586217]

Bates D, Maechler M, Bolker B, Walker S, Haubo R, & Christensen B, (2016). *CRAN Repos* Package lme4: Linear mixed-effects models using Eigen and S4. 1–113

Baus C, Carreiras M, & Emmorey K (2013). When does iconicity in sign language matter? Language and Cognitive Processes, 28(3), 261–271. 10.1080/01690965.2011.620374 [PubMed: 23543899]

Berk S, & Lillo-Martin D (2012). The two-word stage: Motivated by linguistic or cognitive constraints? Cognitive Psychology, 65(1), 118–140. 10.1016/j.cogpsych.2012.02.002 [PubMed: 22475876]

Besner D, & Coltheart M (1979). Ideographic and alphabetic processing in skilled reading of English. Neuropsychologia, 17(5), 467–472. 10.1016/0028-3932(79)90053-8 [PubMed: 514483]

Bosworth RG, & Emmorey K (2010). Effects of iconicity and semantic relatedness on Lexical access in American sign language. Journal of Experimental Psychology: Learning Memory and Cognition, 36(6), 1573–1581. 10.1037/a0020934 [PubMed: 20919784]

Boudreault P, & Mayberry RI (2006). Grammatical processing in American sign language: Age of first-language acquisition effects in relation to syntactic structure. Language and Cognitive Processes, 21(5), 608–635. 10.1080/01690960500139363

Brysbaert M, Mandera P, & Keuleers E (2017). The word frequency effect in word processing: An updated review Current Directions in Psychological Science. 27(1), 45–50. 10.1177/0963721417727521

Bugg JM, DeLosh EL, Davalos DB, & Davis HP (2007). Age differences in stroop interference: Contributions of general slowing and task-specific deficits. Aging, Neuropsychology, and Cognition, 14(2), 155–167. 10.1080/138255891007065

Bull R, Blatto-Vallee G, & Fabich M (2006). Subitizing, magnitude representation, and magnitude retrieval in deaf and hearing adults. Journal of Deaf Studies and Deaf Education, 11(3), 289–302. 10.1093/deafed/enj038 [PubMed: 16601145]

Bull R, Marschark M, & Blatto-Vallee G (2005). SNARC hunting: Examining number representation in deaf students. Learning and Individual Differences, 15(3), 223–236. 10.1016/j.lindif.2005.01.004

Campbell JID, & Alberts NM (2009). Operation-specific effects of numerical surface form on arithmetic strategy. Journal of Experimental Psychology: Learning Memory and Cognition, 35(4), 999–1011. 10.1037/a0015829 [PubMed: 19586266]

Campbell JID, & Epp LJ (2004). An encoding-complex approach to numerical cognition in Chinese-English bilinguals. Canadian Journal of Experimental Psychology, 58(4), 229–244. 10.1037/h0087447 [PubMed: 15648727]

Campbell JID, & Fugelsang J (2001). Strategy choice for arithmetic verification: Effects of numerical surface form. Cognition, 80(3), B21–B30. 10.1016/S0010-0277(01)00115-9 [PubMed: 11274988]

Carey S (2009). Where our number concepts come from. The Journal of Philosophy, 106(4), 220–254. http://www.ncbi.nlm.nih.gov/pubmed/23136450 [PubMed: 23136450]

Cheng Q, & Mayberry RI (2019). Acquiring a first language in adolescence: The case of basic word order in American sign language. Journal of Child Language, 46(2), 214–240. 10.1017/S0305000918000417 [PubMed: 30326985]

Cheng Q, & Mayberry RI (2021). When event knowledge overrides word order in sentence comprehension: Learning a first language after childhood. Developmental Science 24(5), e13073. 10.1111/desc.13073 [PubMed: 33296520]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Cheng Q, Roth A, Halgren E, & Mayberry RI (2019). Effects of early language deprivation on brain connectivity: Language pathways in deaf native and late first-language learners of American sign language. Frontiers in Human Neuroscience, 13, 320. 10.3389/fnhum.2019.00320 [PubMed: 31607879]

Chinello A, de Hevia MD, Geraci C, & Girelli L (2012). Finding the spatial-numerical association of response codes (SNARC) in signed numbers: Notational effects in accessing number representation. Functional Neurology, 27(3), 177–185. http://www.ncbi.nlm.nih.gov/pubmed/23402679 [PubMed: 23402679]

Chrisomalis S (2019). A cognitive typology for numerical notation. Cambridge Archaeological Journal. 14(1), 37–52. 10.1017/S0959774304000034

Chrisomalis S (2020). *Reckonings:* Numerals, cognition, and history. MIT Press.

Cohen DJ (2009). Integers do not automatically activate their quantity representation. Psychonomic Bulletin and Review, 16(2), 332–336. 10.3758/PBR.16.2.332 [PubMed: 19293103]

Cohen Kadosh R, Cohen Kadosh K, Kaas A, Henik A, & Goebel R (2007). Notation-dependent and -independent representations of numbers in the parietal lobes. Neuron, 53(2), 307–314. 10.1016/j.neuron.2006.12.025 [PubMed: 17224410]

Cohen Kadosh R, Gevers W, & Notebaert W (2011). Sequential analysis of the numerical stroop effect reveals response suppression. Journal of Experimental Psychology: Learning Memory and Cognition, 37(5), 1243–1249. 10.1037/a0023550 [PubMed: 21500951]

Cohen Kadosh R, Henik A, & Rubinsten O (2008). Are Arabic and verbal numbers processed in different ways? Journal of Experimental Psychology: Learning, Memory, and Cognition. 34(6), 1377. 10.1037/a0013413 [PubMed: 18980402]

Cohen Kadosh R, & Walsh V (2009). Numerical representation in the parietal lobes: Abstract or not abstract? Behavioral and Brain Sciences, 32(3–4), 313–328. 10.1017/S0140525X09990938 [PubMed: 19712504]

Coppola M, Spaepen E, & Goldin-Meadow S (2013). Communicating about quantity without a language model: Number devices in homesign grammar. Cognitive Psychology, 67(1–2), 1–25. 10.1016/j.cogpsych.2013.05.003 [PubMed: 23872365]

Dehaene S, Bossini S, & Giraux P (1993). The mental representation of parity and number magnitude. Journal of Experimental Psychology: General, 122(3), 371–396. 10.1037/0096-3445.122.3.371

Dehaene S, Dehaene-Lambertz G, & Cohen L (1998). Abstract representations of numbers in the animal and human brain. Trends in Neurosciences, 21(8), 355–361. 10.1016/S0166-2236(98)01263-6 [PubMed: 9720604]

Dehaene S, & Mehler J (1992). Cross-linguistic regularities in the frequency of number words. Cognition, 43(1), 1–29. 10.1016/0010-0277(92)90030-L [PubMed: 1591901]

DeRight J, & Jorgensen RS (2015). I just want my research credit: Frequency of suboptimal effort in a non-clinical healthy undergraduate sample. The Clinical Neuropsychologist, 29(1), 101–117. 10.1080/13854046.2014.989267 [PubMed: 25494327]

Epstein KI, Hillegeist EG, & American JG (1994). Number processing in deaf college students. American Annals of the Deaf, 139(3), 336–347. 10.1353/aad.2012.0321 [PubMed: 7992779]

Ferjan Ramirez N, Leonard MK, Davenport TS, Torres C, Halgren E, & Mayberry RI (2016). Neural language processing in adolescent first-language learners: Longitudinal case studies in American sign language. Cerebral Cortex, 26(3), 1015–1026. 10.1093/cercor/bhu273 [PubMed: 25410427]

Ferjan Ramirez N, Leonard MK, Halgren E, & Mayberry RI (2013). The neural correlates of childhood linguistic isolation. In Baiz RHS & Goldman N (Eds.), 37th Boston University Conference *on language development* (pp. 110–121). Cascadilla Press.

Ferjan Ramirez N, Lieberman AM, & Mayberry RI (2013). The initial stages of first-language acquisition begun in adolescence: When late looks early. Journal of Child Language, 40(2), 391–414. 10.1017/S0305000911000535 [PubMed: 22261245]

Fias W (2001). Two routes for the processing of verbal numbers: Evidence from the SNARC effect. Psychological Research, 65(4), 250–259. 10.1007/s004260100065 [PubMed: 11789429]

Fischer MH, Shaki S, & Cruise A (2009). It takes just one word to quash a SNARC. Experimental Psychology, 56(5), 361–366. 10.1027/1618-3169.56.5.361 [PubMed: 19447752]

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

Flaherty M, & Senghas A (2011). Numerosity and number signs in deaf Nicaraguan adults. Cognition, 121(3), 427–436. 10.1016/j.cognition.2011.07.007 [PubMed: 21899832]

Fromkin V, Krashen S, Curtiss S, Rigler D, & Rigler M (1974). The development of language in genie: A case of language acquisition beyond the "critical period. Brain and Language, 1(1), 81–107. 10.1016/0093-934X(74)90027-3

Gebuis T, Cohen Kadosh R, De Haan E, & Henik A (2009). Automatic quantity processing in 5-year olds and adults. Cognitive Processing, 10(2), 133–142. 10.1007/s10339-008-0219-x [PubMed: 18607652]

Gelman R, & Butterworth B (2005). Number and language: How are they related? Trends in Cognitive Sciences, 9(1), 6–10. 10.1016/j.tics.2004.11.004 [PubMed: 15639434]

Girelli L, Lucangeli D, & Butterworth B (2000). The development of automaticity in accessing number magnitude. Journal of Experimental Child Psychology, 76(2), 104–122. 10.1006/jecp.2000.2564 [PubMed: 10788305]

Gottardis L, Nunes T, & Lunt I (2011). A synthesis of research on deaf and hearing children's mathematical achievement. Deafness and Education International, 13(3), 131–150. 10.1179/1557069X11Y.0000000006

Hall ML, & Dills S (2020). The limits of "communication mode" as a construct. The Journal of Deaf Studies and Deaf Education, 25(4), 383–397. 10.1093/deafed/enaa009 [PubMed: 32432678]

Heine A, Tamm S, De Smedt B, Schneider M, Thaler V, Torbeyns J, Stern E, Verschaffel L, & Jacobs A (2010). The numerical stroop effect in primary school children: A comparison of low, normal, and high achievers. Child Neuropsychology, 16(5), 461–477. 10.1080/09297041003689780 [PubMed: 20437281]

Henik A, & Tzelgov J (1982). Is three greater than five: The relation between physical and semantic size in comparison tasks. Memory & Cognition, 10(4), 389–395. 10.3758/BF03202431 [PubMed: 7132716]

Henner J, Hoffmesiter R, Fish S, Rosenburg P, & DiDonna D (2015). Bilingual instruction works even for deaf children of hearing parents. American Educational Research Association. https://s3.amazonaws.com/academia.edu.documents/52575238/Bilingual_Instruction_Works_Even_for_Dea.pdf?AWSAccessKeyId=AKIAIWOWYYGZ2Y53UL3A&Expires=1530665916&Signature=%2B4l4VdyPZle5P8gi5aIutzWge8c%3D&response-content-disposition=inline%3Bfilename%3DBil

Henner J, Pagliaro C, Sullivan S, & Hoffmeister R (2021). Counting differently: Assessing mathematics achievement in signing deaf and hard of hearing children through a unique lens. American Annals of the Deaf, 166(3), 318–341. 10.1353/aad.2021.0023 [PubMed: 34719519]

Hermans D, Knoors H, Ormel E, & Verhoeven L (2008). The relationship between the reading and signing skills of deaf children in bilingual education programs. Journal of Deaf Studies and Deaf Education, 13(4), 518–530. 10.1093/deafed/enn009 [PubMed: 18400787]

Holmes J, & Adams JW (2006). Working memory and children's mathematical skills: Implications for mathematical development and mathematics curricula. Educational Psychology, 26(3), 339–366. 10.1080/01443410500341056

Hrastinski I, & Wilbur RB (2016). Academic achievement of deaf and hard-of-hearing students in an ASL/English bilingual program. Journal of Deaf Studies and Deaf Education, 21(2), 156–170. 10.1093/deafed/env072 [PubMed: 26864688]

Humphries T, Kushalnagar P, Mathur G, Napoli DJ, Padden C, Rathmann C, & Smith S (2016). Avoiding linguistic neglect of deaf children. Social Service Review, 90(4), 589–619. 10.1086/689543

Iversen W, Nuerk H-C, Jäger L, & Willmes K (2006). The influence of an external symbol system on number parity representation, or What's odd about 6? Psychonomic Bulletin & Review, 13(4), 730–736. 10.3758/BF03193988 [PubMed: 17201377]

Iversen W, Nuerk HC, & Willmes K (2004). Do signers think differently? The processing of number parity in deaf participants. Cortex; a Journal Devoted to the Study of the Nervous System and Behavior, 40(1), 176–178. 10.1016/S0010-9452(08)70940-7 [PubMed: 15174461]

Kaufmann L, Ischebeck A, Weiss E, Koppelstaetter F, Siedentopf C, Vogel SE, Gotwald T, Marksteiner J, & Wood G (2008). An fMRI study of the numerical Stroop task in individuals with and without

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

minimal cognitive impairment. Cortex, 44(9), 1248–1255. 10.1016/j.cortex.2007.11.009 [PubMed: 18761138]

Koluchová J (1972). Severe deprivation in twins: A case study. Journal of Child Psychology and Psychiatry, 13(2), 107–114. 10.1111/j.1469-7610.1972.tb01124.x [PubMed: 5065080]

Koluchová J (1976). The further development of twins after severe and prolonged deprivation: A second report. Journal of Child Psychology and Psychiatry, 17(3), 181–188. 10.1111/j.1469-7610.1976.tb00390.x [PubMed: 945807]

Kritzer KL (2009). Barely started and already left behind: A descriptive analysis of the mathematics ability demonstrated by young deaf children. Journal of Deaf Studies and Deaf Education, 14(4), 409–421. 10.1093/deafed/enp015 [PubMed: 19596725]

Lange CM, Lane-Outlaw S, Lange WE, & Sherwood DL (2013). American sign language/English Bilingual model: A longitudinal study of academic growth. Journal of Deaf Studies and Deaf Education, 18(4), 532–544. 10.1093/deafed/ent027 [PubMed: 23741074]

Liu X, Wang H, Corbly CR, Zhang J, & Joseph JE (2006). The involvement of the inferior parietal cortex in the numerical stroop effect and the distance effect in a two-digit number comparison task. Journal of Cognitive Neuroscience, 18(9), 1518–1530. 10.1162/jocn.2006.18.9.1518 [PubMed: 16989552]

Marshall C, Jones A, Denmark T, Mason K, Atkinson J, Botting N, & Morgan G (2015). Deaf children’s non-verbal working memory is impacted by their language experience. Frontiers in Psychology, 6((MAY), 527). 10.3389/fpsyg.2015.00527 [PubMed: 25999875]

Mathôt S, Schreij D, & Theeuwes J (2012, June). OpenSesame: An open-source, graphical experiment builder for the social sciences. Behavior Research Methods, 44(2), 314–324. 10.3758/s13428-011-0168-7 [PubMed: 22083660]

Mayberry RI, Cheng Q, Hatrak M, & Ilkbasaran D (2017). Late L1 learners acquire simple but not syntactically complex structures. International Association for the Study of Child Language. 2017 Lyon.

Mayberry RI, Davenport T, Roth A, & Halgren E (2018). Neurolinguistic processing when the brain matures without language. Cortex, 99, 390–403. 10.1016/j.cortex.2017.12.011 [PubMed: 29406150]

Mayberry RI, Hatrak M, Ilkbasaran D, Cheng Q, & Hall ML (in prep.). Maturational Constraints on language development: Evidence for a closing of the critical period for language.

Mayberry RI, & Kluender R (2018). Rethinking the critical period for language: New insights into an old question from American sign language. Bilingualism: Language *and Cognition*, 21(5), 886–905. 10.1017/S1366728917000724 [PubMed: 30643489]

Mayberry RI, & Lock E (2003). Age constraints on first versus second language acquisition: Evidence for linguistic plasticity and epigenesis. Brain and Language, 87(3), 369–384. 10.1016/S0093-934X(03)00137-8 [PubMed: 14642540]

Metcalfe AWS, & Campbell JID (2007). The role of cue familiarity in adults’ strategy choices for simple addition. European Journal of Cognitive Psychology, 19(3), 356–373. 10.1080/09541440600872001

Mitchell RE, & Karchmer MA (2004). Chasing the mythical ten percent: Parental hearing status of deaf and hard of hearing students in the United States. Sign Language Studies, 4(2), 47. 10.1353/sls.2004.0005

Newport EL (1990). Maturational Constraints on Language Learning. COGNlTiVE SCIENCE, 14(1), 11–28. https://onlinelibrary.wiley.com/doi/pdf/10.1207/s15516709cog1401_2

Pagliaro CM, & Kritzer KL (2013). The math gap: A description of the mathematics performance of preschool-aged deaf/hard-of-hearing children. Journal of Deaf Studies and Deaf Education, 18(2), 139–160. 10.1093/deafed/ens070 [PubMed: 23307889]

Pansky A, & Algom D (2002). Comparative judgment of numerosity and numerical magnitude: Attention preempts automaticity. Journal of Experimental Psychology. Learning, Memory, and Cognition, 28(2), 259–274. http://www.ncbi.nlm.nih.gov/pubmed/11911383 [PubMed: 11911383]

Pénicaud S, Klein D, Zatorre RJ, Chen J-K, Witcher P, Hyde K, & Mayberry RI (2013). Structural brain changes linked to delayed first language acquisition in congenitally deaf individuals. NeuroImage, 66, 42–49. 10.1016/j.neuroimage.2012.09.076 [PubMed: 23063844]

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

R CORE TEAM. (2016). R: A language and environment for statistical computing. 2018. R Foundation for Statistical Computing, Vienna, Austri.

Razpurker-Apfeld I, & Koriat A (2006). Flexible mental processes in numerical size judgments: The case of Hebrew letters that are used to convey numbers. Psychonomic Bulletin & Review, 13(1), 78–83. 10.3758/BF03193816 [PubMed: 16724772]

Rousselle L, & Noël MP (2007). Basic numerical skills in children with mathematics learning disabilities: A comparison of symbolic vs non-symbolic number magnitude processing. Cognition, 102(3), 361–395. 10.1016/j.cognition.2006.01.005 [PubMed: 16488405]

Rubinsten O, & Henik A (2005). Automatic activation of internal magnitudes: A study of developmental dyscalculia. Neuropsychology, 19(5), 641–648. 10.1037/0894-4105.19.5.641 [PubMed: 16187882]

Rubinsten O, Henik A, Berger A, & Shahar-Shalev S (2002). The development of internal representations of magnitude and their association with Arabic numerals. Journal of Experimental Child Psychology, 81(1), 74–92. 10.1006/jecp.2001.2645 [PubMed: 11741375]

Schwarz W, & Heinze HJ (1998). On the interaction of numerical and size information in digit comparison: A behavioral and event-related potential study. Neuropsychologia, 36(11), 1167–1179. 10.1016/S0028-3932(98)00001-3 [PubMed: 9842762]

Sehyr ZS, Caselli N, Cohen-Goldberg AM, & Emmorey K (2021). The ASL-LEX 2.0 project: A database of Lexical and phonological properties for 2,723 signs in American sign language. The Journal of Deaf Studies and Deaf Education, 26(2), 263–277. 10.1093/deafed/enaa038 [PubMed: 33598676]

Shaki S, & Fischer MH (2008). Reading space into numbers – A cross-linguistic comparison of the SNARC effect. Cognition, 108(2), 590–599. 10.1016/j.cognition.2008.04.001 [PubMed: 18514179]

Shaki S, Fischer MH, Petrusic WM, & Shaki S (2009). Reading habits for both words and numbers contribute to the SNARC effect. Psychonomic Bulletin & Review, 16(2), 328–331. 10.3758/PBR.16.2.328 [PubMed: 19293102]

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

Spaepen E, Coppola M, Spelke ES, Carey SE, & Goldin-Meadow S (2011). Number without a language model. Proceedings of the National Academy of Sciences, 108(8), 3163–3168. 10.1073/pnas.1015975108

Spelke ES (2017). Language learning and development core knowledge, language, and number. Language Learning and Development, 13(2), 147–170. 10.1080/15475441.2016.1263572

Szucs D, & Soltész F (2007). Event-related potentials dissociate facilitation and interference effects in the numerical Stroop paradigm. Neuropsychologia, 45(14), 3190–3202. 10.1016/j.neuropsychologia.2007.06.013 [PubMed: 17675108]

Takahashi A, & Green D (1983). Numerical judgments with Kanji and Kana. Neuropsychologia, 21(3), 259–263. 10.1016/0028-3932(83)90042-8 [PubMed: 6877579]

Taub SF (2001). Language from the body: Iconicity and metaphor in American sign language (Cambridge University Press.).

Traxler CB (2000). The Stanford achievement test, 9th edition: National norming and performance standards for deaf and hard-of-hearing students. Journal of Deaf Studies and Deaf Education, 5(4), 337–348. 10.1093/deafed/5.4.337 [PubMed: 15454499]

Tzelgov J, Meyer J, & Henik A (1992). Automatic and intentional processing of numerical information. Journal of Experimental Psychology: Learning, Memory, and Cognition, 18(1), 166–179. 10.1037/0278-7393.18.1.166

Vaid J (1985). Numerical size comparisons in a phonologically transparent script. *Perception &* Psychophysics, 37(6), 592–595. 10.3758/BF03204927

Vaid J, & Corina D (1989). Visual field asymmetries in numerical size comparisons of digits, words, and signs. Brain and Language, 36(1), 117–126. http://www.ncbi.nlm.nih.gov/pubmed/2917283 [PubMed: 2917283]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Verhaeghen P, & De Meersman L (1998). Aging and the Stroop effect: A meta-analysis. Psychology and Aging, 13(1), 120–126. 10.1037/0882-7974.13.1.120 [PubMed: 9533194]

West R, & Baylis GC (1998). Effects of increased response dominance and contextual disintegration on the Stroop interference effect in older adults. Psychology and Aging, 13(2), 206–217. http://www.ncbi.nlm.nih.gov/pubmed/9640582 [PubMed: 9640582]

Whelan R (2008). Effective analysis of reaction time data. The Psychological Record, 58(3), 475–482. 10.1007/BF03395630

White SLJ, Szucs D, & Soltész F (2012). Symbolic number: The integration of magnitude and spatial representations in children aged 6 to 8years. Frontiers in Psychology, 3(JAN) 392.10.3389/fpsyg.2011.00392 [PubMed: 23091464]

Wood G, Willmes K, Nuerk H-C, & Fischer MH (2008). On the cognitive link between space and number: A meta-analysis of the SNARC effect. Psychology Science Quarterly, 50(4), 489. https://bit.ly/35HO2af

Wood D, Wood H, & Howarth P (1983). Mathematical abilities of deaf school-leavers. British Journal of Developmental Psychology, 1(1), 67–73. 10.1111/j.2044-835X.1983.tb00544.x

Wood HA, Wood DJ, Kingsmill MC, French JR, & Howarth SP (1984). The mathematical achievements of deaf children from different educational environments. The British Journal of Educational Psychology, 54 (Pt 3)(3), 254–264. 10.1111/j.2044-8279.1984.tb02589.x [PubMed: 6508996]

Zarfaty Y, Nunes T, & Bryant P (2004). The performance of young deaf children in spatial and temporal number tasks. Journal of Deaf Studies and Deaf Education, 9(3), 315–326. 10.1093/deafed/enh034 [PubMed: 15304434]

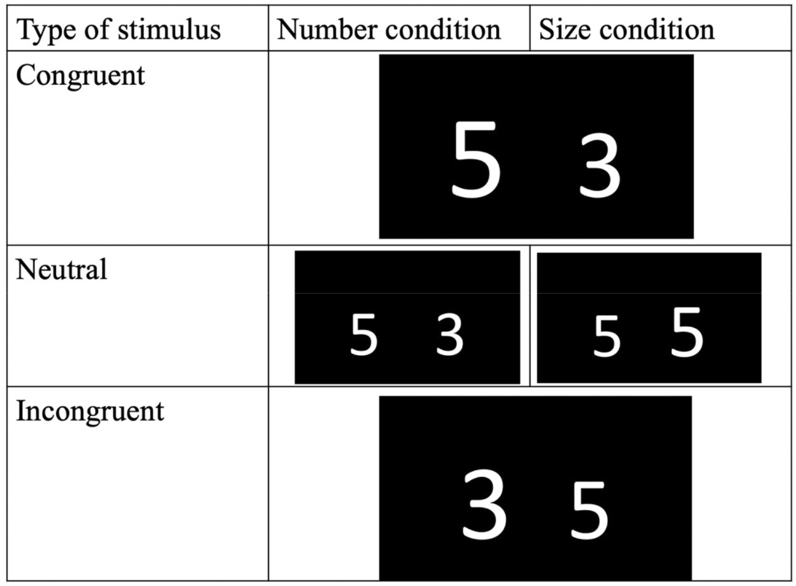


Figure 1.
Example of stimuli in Number Stroop Test.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Author Manuscript

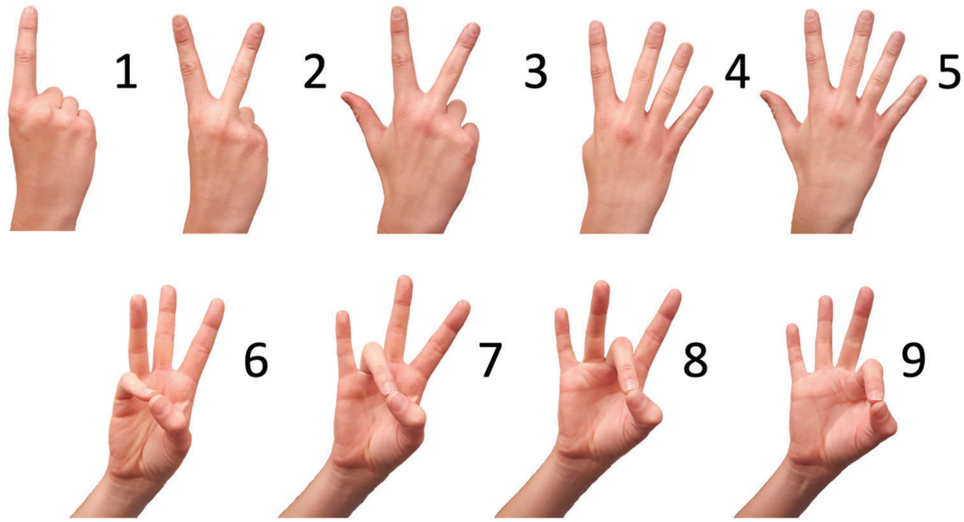


Figure 2.
ASL number signs ONE – NINE.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

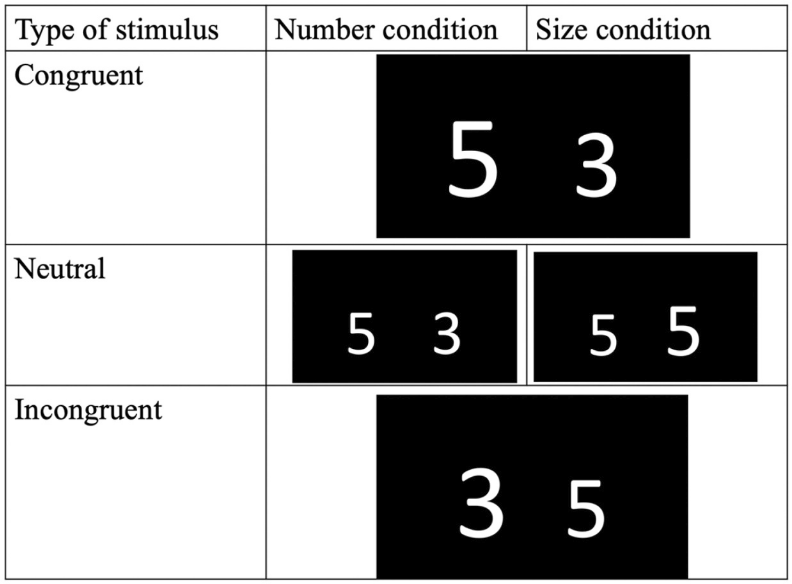


Figure 3.
Examples of stimuli for Arabic digit block.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

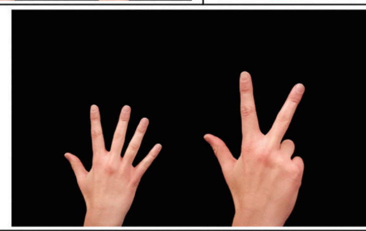
Type of stimulus	Number condition	Size condition
Congruent		
Neutral		
Incongruent		

Figure 4.
Examples of stimuli for ASL block.

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

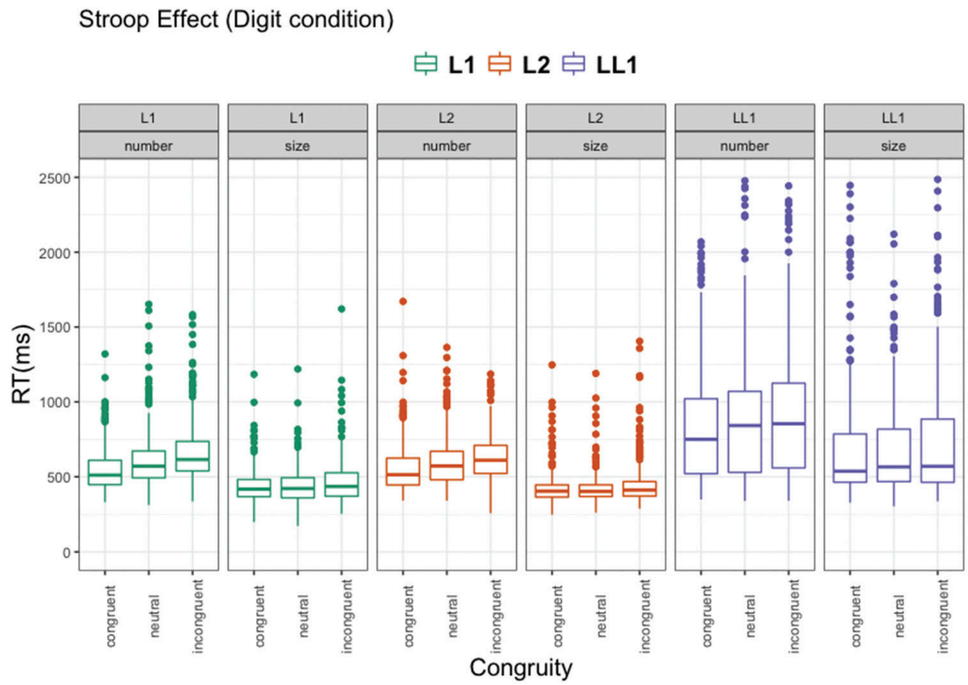


Figure 5.

Response times for the trials with Arabic digits. The head of the facet and the color indicate a group of participants (L1, L2, or LL1) and the condition (type or number). The top of the box plot shows the higher quartile (75%), the bar shows the median (50%), and the bottom of the box shows the lower quartile (25%); the dots show outliers outside the 1.5 interquartile range.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

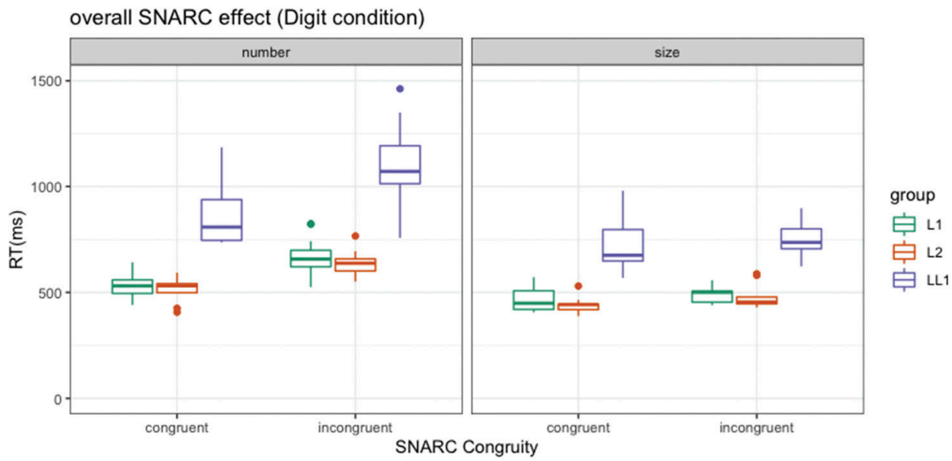


Figure 6.

Overall SNARC effect (when size and number information is both congruent or incongruent) with Arabic digits. The colors indicate a group of participants (L1, L2, or LL1) and the columns show SNARC congruity (congruent or incongruent). The top of the box plot shows the higher quartile (75%), the bar shows the median (50%), and the bottom of the box shows the lower quartile (25%); the dots show outliers outside the 1.5 interquartile range.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

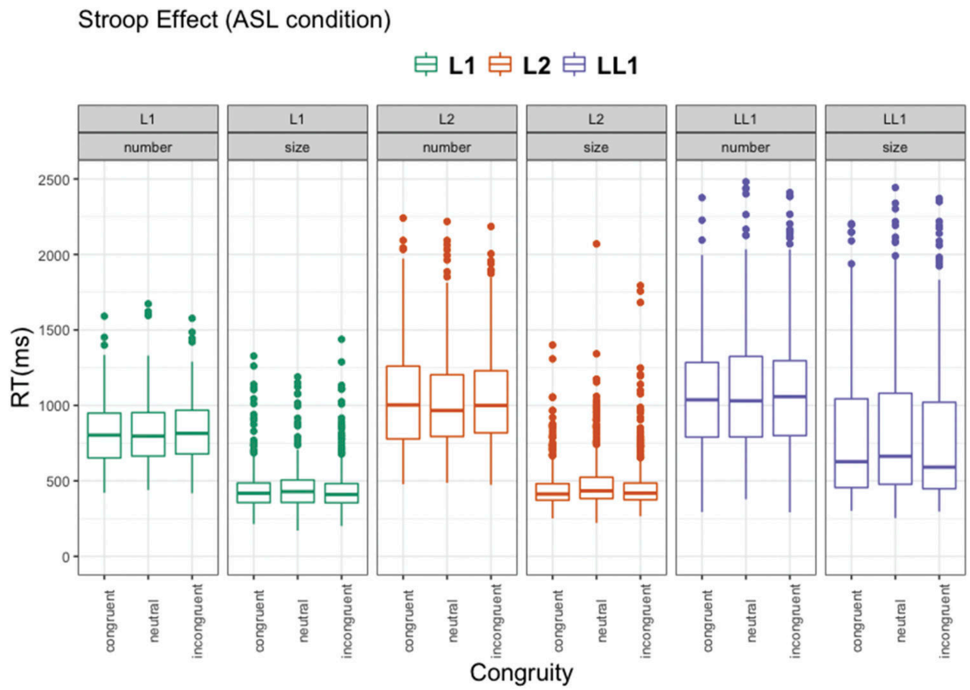


Figure 7.

Response times for trials with ASL signs. The head of the facet and the color indicate a group of participants (L1, L2, LL1) and the condition (size and number). The top of the box plot shows the higher quartile (75%), the bar shows the median (50%), and the bottom of the box shows the lower quartile (25%); the dots show outliers outside of 1.5 interquartile range.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1.

Possible outcomes of Arabic digit task.

Number Stroop Effect with Arabic Digits	Possible Interpretation
Found in all groups; no differences	Age of acquisition does not affect automatic magnitude processing
Found in all groups, but there are differences in late first language learners	Age of acquisition affects magnitude processing, but automatic processing still can be achieved despite incomplete early input (i.e., only digits)
Found in all groups but late learners	Age of acquisition affects magnitude processing, without early language exposure automatic magnitude processing is not achieved
Found only in hearing second language learners	Something other than language deprivation affects automatic magnitude processing in deaf participants

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

Table 2.

Possible outcomes of ASL task.

Number Stroop Effect with ASL signs	Possible Interpretation
Found in all groups	ASL number signs activate magnitude information in the same way as Arabic digits, supporting the common number representation hypothesis
Not found in all groups	ASL number signs activate magnitude in a different way from Arabic digits, supporting the format-specific activation hypothesis
Late first language learners differ from other groups	Age of acquisition rather than years of exposure influences automatic magnitude processing with number signs
Hearing signers differ from other groups	Years of exposure rather than age of acquisition influence automatic magnitude processing with number signs
Early first language learners differ from other groups	Both years of exposure and age of acquisition influence automatic magnitude processing with number signs

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 3.

Possible outcomes of Arabic digit task.

Arabic digits								
Size mean RT (SD)/ Accuracy (SD)					**Number mean RT (SD)/ Accuracy (SD)**			
	congruent	**neutral**	**incongruent**	**overall mean**	**congruent**	**neutral**	**incongruent**	**overall mean**
LI (early first language learners)	438 (112)/0.99 (0.009)	437 (106)/ 0.90(0.02)	473 (153)/ 0.97(0.03)	450 (127)/ 0.95(0.04)	549 (146)/0.99 (0.009)	613 (188)/0.97(0.03)	678 (208)/0.92 (0.11)	612 (189)/ 0.96(0.07)
L2 (second language learners)	424 (104)/ 0.98(0.01)	422 (96)/ 0.90(0.01)	450 (138)/ 0.96(0.03)	432 (116)/ 0.95(0.04)	558 (162)/0.98 (0.01)	601 (163)/0.96 (0.02)	652 (156)/0.89 (0.07)	602 (165)/ 0.95(0.05)
LL1 (late first language learners)	674 (378)/ 0.94(0.15)	642 (279)/ 0.85(0.19)	698 (360)/ 0.89(0.23)	672 (343)/ 0.89(0.19)	823 (386)/ 0.99(0.018)	931 (494)/0.91(0.17)	942 (458)/ 0.72(0.023)	893 (448)/ 0.88(0.23)

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 4.

Result summary for number Stroop effect with Arabic digits.

	Response time		
Predictors	***Estimates***	***CI***	***P***
(Intercept)	604.64	492.92–724.2	**<.001**
congruity [congruent]	−31.98	58.7–7.95	**.014**
congruity [incongruent]	49.94	23.73–75.34	**<.001**
AoA [L2]	−10.47	−166.37–162.63	.902
AoA [LL1]	306.88	126.42–469.98	**<.001**
condition [size]	−158.63	−176.34 – −139.53	**<.001**
congruity [congruent] * AoA [L2]	11.69	−16.76–41.21	.441
congruity [incongruent] * AoA [L2]	−12.53	−41.12–18.07	.415
congruity [congruent] * AoA [LL1]	−7.22	−37.20–21.04	.641
congruity [incongruent] * AoA [LL1]	−12.05	−42.94–19.57	.450
AoA [L2] * condition [size]	−6.73	−29.67–18.03	.589
AoA [LL1] * condition [size]	−55.63	−81.72 – −30.19	**<.001**
Random Effects			
σ^2		37529.17	
$\tau_{00\ \mathrm{stimulus}}$		384.04	
$\tau_{00\ \mathrm{block\ order:subject}}$		33433.23	
ICC		0.47	
N_{stimulus}		48	
$N_{\mathrm{block\ order}}$		2	
N_{subject}		29	
Observations		5728	
Marginal R^2/Conditional R^2		0.262/0.612	

Results summary for Number Stroop Effect with Arabic digits.

Age of Acquisition (AoA) groups: L1 (early first language learners), L2 (second language learners), LL1 (late first language learners).

Conditions: size and number.

Congruity levels: congruent, neutral, incongruent.

Reference categories: congruity = neutral, AoA = L1, condition = number.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 5.

Mean RT (SD) as a function of congruity.

ASL signs.	Size mean RT (SD)/Accuracy (SD)				Number mean RT (SD)/ Accuracy (SD)			
	congruent	**neutral**	**incongruent**	**overall mean**	**congruent**	**neutral**	**incongruent**	**overall mean**
LI (early first language learners)	445 (145)/0.98 (0.02)	452 (146)/0.99 (0.02)	452 (164)/0.98 (0.01)	450 (152)/0.98 (0.02)	808 (204)/0.95 (0.02)	832 (219)/0.95 (0.03)	841 (197)/0.92 (0.07)	827 (207)/0.94 (0.05)
L2 (second language learners)	455 (147)/0.98 (0.01)	498 (192)/0.97 (0.02)	473 (194)/0.98 (0.01)	476 (180)/0.98 (0.02)	1049 (335)/0.95 (0.05)	1030 (330)/0.94 (0.05)	1063 (329)/0.95 (0.04)	1047 (331)/ 0.95 (0.04)
LL1 (late first language learners)	802 (444)/0.97 (0.02)	846 (481)/ 0.97(0.02)	755 (445)/0.92 (0.15)	802 (458)/0.96 (0.09)	1079 (412)/0.98 (0.01)	1130 (446)/0.91 (0.17)	1134 (391)/0.78 (0.39)	1112 (418)/ 0.89 (0.24)

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 6.

Result summary for number Stroop effect with ASL sign.

	response_time		
Predictors	***Estimates***	***CI***	***P***
(Intercept)	836.13	690.74–983.89	**<.001**
congruity [congruent]	−14.37	−59.48–30.49	.546
congruity [incongruent]	4.50	−39.32–53.94	.850
AoA [L2]	221.80	33.48–425.65	**.018**
AoA [LL1]	302.93	118.53–494.29	**.001**
condition [size]	−385.58	−410.99 – −361.42	**<.001**
congruity [congruent] * AoA [L2]	−8.61	−49.51–34.21	.670
congruity [incongruent] * AoA [L2]	−4.92	−41.41–36.68	.808
congruity [congruent] * AoA [LL1]	−40.26	−78.68 – −2.58	**.044**
congruity [incongruent] * AoA [LL1]	−39.51	−77.82 – −2.41	.052
AoA [L2] * condition [size]	−190.91	−226.04 – −160.40	**<.001**
AoA [LL1] * condition [size]	84.03	49.44–115.34	**<.001**
Random Effects			
σ^2		60460.57	
$\tau_{00\ stimulus}$		2875.89	
$\tau_{00\ subject\_parity:subject}$		40513.91	
ICC		0.42	
$N_{stimulus}$		48	
$N_{subject\_parity}$		2	
$N_{subject}$		29	
Observations		5379	
Marginal R^2/Conditional R^2		0.387/0.643	

Age of Acquisition (AoA) groups: L1 (early first language learners), L2 (second language learners), LL1 (late first language learners).

Conditions: size and number.

Congruity levels: congruent, neutral, incongruent.

Reference categories: congruity = neutral, AoA = L1, condition = number.

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

Table 7.

Result summary for Iconicity Effect with ASL signs.

	rt			rt		
Predictors	***Estimates***	***CI***	***P***	***Estimates***	***CI***	***P***
(Intercept)	932.44	870.77–994.11	**<.001**	950.29	879.48–1021.11	**<.001**
congruity_stroop [neutral]	61.27	−0.40–122.94	.051	65.55	7.06–124.04	**.028**
congruity_stroop [incongruent]	33.29	−28.38–94.96	.288	34.84	−23.65–93.33	.242
AoA[L2]	192.56	130.89–254.22	**<.001**	192.56	134.12–250.99	**<.001**
AoA [LL1]	455.17	393.51–516.84	**<.001**	455.17	396.74–513.61	**<.001**
condition [size]	546.62	596.98–496.27	**<.001**	−544.80	−592.61–496.99	**<.001**
iconicity [−]				37.20	−23.79–98.19	.231
iconicity [+]				−102.79	−165.62 – −39.95	**.001**
Observations		216			216	
R^2/R^2 adjusted		0.763/0.757			0.789/0.782	

Results summary for Iconicity Effect with ASL signs.

Age of Acquisition (AoA) groups: L1 (early first language learners), L2 (second language learners), LL1 (late first language learners). Iconicity: iconic. [+], non-iconic [−], mix.

Stroop congruity levels: congruent, neutral, incongruent.

Reference categories: congruity = congruent, AoA = L1, condition = number, iconicity = mix.