Generalized slowing rather than inhibition is associated with language outcomes in both late talkers and children with typical early development

Anna Kautto[1]

Eira Jansson-Verkasalo[1]

Elina Mainela-Arnold[1, 2]

[1]Department of Psychology and Speech-Language Pathology, University of Turku, Finland

[2]Department of Speech-Language Pathology, University of Toronto, ON, Canada

Conflicts of interest: The authors report no conflicts of interest

Funding statement: This research was financially supported by the University of Turku Graduate School wages awarded to the first author and an anonymous endowed fund to the University of Turku Speech-Language Pathology.

Contact information:

Anna Kautto

anna.kautto@utu.fi

Abstract

Purpose

While most of the children who are identified as late talkers at the age of two catch up their peers before school-age, some continue to have language difficulties and will later be identified as having developmental language disorder. Our understanding of which children catch up and which do not is limited. The aim of the current study was to find out if inhibition is associated with late talker outcomes at school-age.

Method

We recruited 73 school-aged children (ages 7–10 years) with a history of late talking (n = 38) or typical development (n = 35). Children completed measures of language skills and a flanker task to measure inhibition. School age language outcome was measured as a continuous variable.

Results

Our analyses did not reveal associations between inhibition and school-aged language index or history of late talking. However, stronger school-age language skills were associated with shorter overall response times on the flanker task, both in congruent and incongruent trials. This effect was not modulated by history of late talking, suggesting that a relationship between general response times and language development is similar in both children with typical early language development and late talkers.

Conclusions

Inhibition is not related to late talker language outcomes. However, children with better language outcomes had shorter general response times. We interpret this to reflect differences in

general processing speed, suggesting that processing speed holds promise for predicting school-age language outcomes in both late talkers and children with typical early development.

The term “late talker” (LT) is used to refer to toddlers with late onset of first words and unusually small vocabulary size without any other developmental or hearing impairments which would otherwise explain the reduced vocabulary (Fisher, 2017; Rescorla, 2011). Research has shown that late talkers are at an increased risk for persistent language difficulties (Fisher, 2017). However, late talking itself has relatively low predictive value for persistent language difficulties (Rescorla, 2011) varying from 14.4 - 17.6 % (Westerlund et al., 2006) to 47 - 73 % (Feldman et al., 2005), depending on age points and criteria used in defining late talking and language disorder. Many late talkers catch up with their peers in language skills by school age. These children are often referred to as late bloomers. However, we still lack sufficient knowledge about which late talkers will have long-term restrictions in language abilities (see Moyle, Stokes, & Klee, 2011).

In this study, we examined if presence of limitations in inhibition co-occur with poor language abilities in school-aged children with and without a history of late talking. Inhibition is a part of executive functions, which are a set of mental processes that are involved in controlling and regulating behaviors (Diamond, 2013). Executive functions are often divided into three main areas: shifting, updating, and inhibition (Lehto et al., 2003; Miyake et al., 2000). Shifting refers to the ability to change between tasks and cognitive flexibility. Updating is closely related to working memory, storing and manipulating information. Inhibition refers to controlling behavior and choosing reactions; in the context of attention especially being able to ignore irrelevant and distracting information while selecting crucial information (Diamond, 2013).

Executive functions develop early in childhood (Marcovitch & Zelazo, 2009). Several theories of language development consider development of executive function to be closely related to language development. For instance, according to Merriman (1999) when acquiring language children learn to attend to features of stimuli in the language learning environment that were valued and relevant in similar contexts earlier. Indeed, children's vocabulary (Chow et al., 2019) and sentence level abilities (Kaushanskaya et al., 2017) have been associated with measures of inhibition skills.

According to Rueda et al. (2004), inhibition skills typically clearly improve from 6 to 7 years of age, but remain stable after that with performance on inhibition tasks reaching adult like levels around 10 years of age. As limitations in inhibition co-occur with poor language abilities (Finneran et al., 2009; Henry et al., 2012; Larson et al., 2020; Marton et al., 2007, 2018; Park et al., 2019), we might be able to use measures of inhibition to identify children at highest risk for persistent language difficulties. This would give us the possibility to cost-effectively focus on prevention and early intervention of language disorders.

Developmental language disorder or specific language impairment

If children over the age of four continue to exhibit language difficulties, the term Developmental Language Disorder (DLD) can be used (Bishop et al., 2017). In DLD, deficits in different areas of language are seen, including deficits in morphosyntactic, semantic or lexical abilities. By the most recent definition, the term DLD can be used when children's language difficulties interfere with daily life, and when factors associated with poor prognosis are present (Bishop et al., 2016). Some of the risk factors associated with poor prognosis include family history of language disorders and being male. These risk factors are statistically associated with DLD, but the causal relationship between risk factors and DLD is unclear (Bishop et al., 2017). As the risk factors cumulate, the likelihood of persistent language difficulties increases.

While differential diagnoses for DLD include autism spectrum disorders, genetic conditions like Down syndrome, cerebral palsy, intellectual disability and hearing loss, DLD can co-occur with some neurodevelopmental difficulties, such as problems in motor skills, social interaction and attention (Bishop et al., 2017). Indeed, attention deficits are commonly observed in children with DLD (see Mueller & Tomblin, 2012). Children with DLD have also been suggested to exhibit difficulties in areas of executive functions, including nonverbal inhibition (Henry et al., 2012). To our knowledge, it is still unknown how deficits in inhibition are related to late talking and late blooming. Inhibition deficits could possibly be one of the cumulative risk factors for DLD. In other words, we postulated that poor inhibition abilities would be associated with lower school age language abilities in children with a history of late talking.

Before the term DLD was launched, the term Specific language impairment (SLI) was commonly used to refer to children with language difficulties without any additional biomedical conditions. The use of the term DLD instead of SLI has rapidly increased after the publication of CATALISE consensus papers concerning preferred terminology (Bishop et al., 2016; Bishop et al., 2017). However, some researchers still advocate using the term SLI. One of the key differences between the definitions of DLD and SLI lies in how the relationship between verbal and nonverbal performance is defined. The definition of DLD does not require a difference between verbal and nonverbal skills and children with low within normal non-verbal IQ are included. In contrast, the SLI definition requires that nonverbal IQ should be within age expectations and stringent non-verbal IQ criteria are often employed (e.g. > 85). Consequently, we cannot directly assume findings from the SLI populations to generalize to DLD or vice versa. Leonard (2020) suggested that when researchers conduct research relevant to DLD, they should also provide results meeting stricter non-verbal IQ criteria, which advice we followed in this study.

Inhibition and DLD

Even though the etiology for DLD is still unknown, several theories concerning the underlying impairment have been proposed. For example, it has been suggested that deficits in language acquisition could be a result of processing capacity limitations (Gathercole & Baddeley, 1990; Miller et al., 2001; Montgomery et al., 2009). According to the limited processing capacity theories, children with DLD have limited capacity for processing linguistic stimuli resulting in incomplete learning. The capacity theories differ from each other in the assumptions they make about the nature of capacity limitations. One way in which children's processing capacity can be limited is limitations in attentional inhibition. To acquire language, one must be able to orient to and keep focus on incoming stimuli while simultaneously suppressing irrelevant information (Ebert & Kohnert, 2011). When a child has difficulties in inhibition, irrelevant information may overload processing and degrade acquiring the crucial part of the input.

Consistent with the hypothesis that limitations in inhibition constrain language development, children with DLD have been shown to differ from their peers in two aspects of attention, inhibition and sustained attention. Inhibition refers to selection of crucial information and ignoring distracting or irrelevant information. Sustained attention refers to the ability to maintain alertness and focus over time (Mirsky et al., 1991). A meta-analysis by Pauls and Archibald (2016) indicated that children with DLD have attentional difficulties, especially in inhibition. After the publication of the meta-analysis, other studies have also investigated the relationship between DLD and inhibition, and suggested that deficits in inhibition are associated with language difficulties (Larson et al., 2020; Marton et al., 2018). Larson et al. (2020) suggest that inhibition predicts morphological comprehension during one-year follow up in school-aged children with DLD.

Inhibition is often measured with flanker (Eriksen & Eriksen, 1974) or Simon (Simon & Wolf, 1963) tasks in which the stimuli occur with distracting or irrelevant information and effects of the distraction are measured. On the flanker tasks, participants are asked to focus on an array of arrows and indicate the direction the center (target) arrow is pointing to as quickly as possible. On the congruent trials, all five arrows point to the same direction. On the incongruent trials, the target (center) arrow point to different direction than the other arrows (distractors). The effect of congruency, or the flanker effect, is the difference in performance between the congruent and incongruent trials. Participants are typically responding slower (and often also less accurately) in incongruent as opposed to congruent trials. In children's version of the task, arrows are often replaced with fish (Figure 1). Ebert et al. (2019) and Park et al. (2019) reported that children with DLD exhibited weaker inhibition than typically developing peers, as reflected by larger flanker effect.

The meta-analysis by Ebert and Kohnert (2011) also provides evidence suggesting difficulties in sustained attention in children with DLD. While these difficulties are present in both auditory and visual domains, effect sizes are larger in studies using auditory and linguistic as opposed to visual stimuli. The authors further point out that many cognitive tasks not intended to measure attention contain a component of sustained attention and poor task performance in children with DLD can at least in part be explained by difficulties in attention (Im-Bolter et al., 2006). Even though there is a body of converging evidence for attention difficulties in DLD (see Ebert & Kohnert, 2011; Pauls & Archibald, 2016), to our knowledge, there are no published studies investigating attention in children with confirmed late talking histories.

Other theories have postulated that the crucial capacity limitation constraining language learning is not limitations in inhibition but processing speed instead. According to generalized slowing hypothesis, slow processing speed across domains restricts language learning in children with DLD (Kail, 1994). It is therefore important to consider the generalized slowing hypothesis when using tasks involving RTs to examine inhibition. The flanker task is designed so that it differentiates between more general slowness and inhibition difficulties by comparing RTs in congruent and incongruent trials. While generalized slowing would be observed as long RTs across both trial types, difficulties in inhibition would be observed in slower performance in incongruent trials as compared to congruent, because the inhibition load is larger in these trials.

The relationship between late talking and DLD

Although the relationship between late talking and DLD is still unclear, investigators have made assumptions about possible shared and distinct underlying mechanisms. According to the categorical approach, late talking and persistent language disorder, DLD, have different etiologies and outcomes (Rescorla, 2009). Research based on categorical approach often focuses on possible genetic causes or clinical markers (such as grammatical skills) of DLD. If the mechanisms underlying late talking and DLD are different, it would be possible to find a feature, a gene, behavioral or other marker that could be used to categorically predict whether a late talking child will have persistent language impairment.

In contrast to the categorical approach, the *dimensional account* (Rescorla, 2009, 2011) states that late talkers and typically developing peers “differ quantitatively on a hypothetical language ability spectrum” (Rescorla, 2011, p.141). These differences would be due to variation in skills subserving language. These subserving skills are hypothesized to be similar to those that have been suggested to underlay DLD. Supporting the dimensional account of late talking, Rescorla reported that most late talkers who score within the normal range on language tests during school age still continue to perform significantly lower than their typically developing peers. Thus, in this study we chose to treat language abilities in children with the history of late talking as a continuum with school age language abilities ranging from severe to mild to no difficulties instead of categorical division of children into to DLD, late blooming and typical language trajectory groups.

Following the dimensional account on late talking assuming that DLD and late talking share similar critical skills subserving language learning, we postulated that limitations in inhibition underlie both late blooming and persistent language difficulties. The difference between persistent language difficulties and late blooming would be that of a degree of limitation in inhibition difficulties. Children with a history of late talking would have lower inhibition skills than children with typical early language development and the inhibition ability would also be associated with language abilities at school age, especially in late talkers.

Current study

Late talkers are known to exhibit an increased risk of DLD. However, predicting late talker outcomes has been proven to be difficult. Given that limitations in inhibition have been reported in children with DLD and those limitation have been hypothesized to underlie the persistent language difficulties these children face, we investigated the associations between inhibition, early late talking status and language skills in school aged children. If such a relationship exists, subsequent research on inhibition skills in children who are late talkers may confirm that inhibition skills can be used in predicting which children will catch up with their peers and which will continue to have persistent language difficulties. We hypothesized that weak performance on an inhibition task would be associated with poor language skills.

Furthermore, based on the dimensional account, we hypothesized that (1) late talkers would exhibit a larger flanker effect as compared to children without LT history and that (2) flanker effect would be larger in late talkers with low school age language abilities as compared to LT and TED peers with higher school-age language index.

Methods

Participants

Participants were recruited from the Southwestern Birth Cohort study (Lagström et al., 2012). This cohort study included a total of 9936 children, 1827 of which participated in subsequent studies. Both the birth cohort study and the current study were approved by the Ethics Committee of the Hospital District of Southwest Finland.

We sent invitation letters to the cohort families with children who met criteria for late talking at 24 or 36 months of age. Late talking (LT) was defined as either (1) expressive vocabulary below the 11th percentile on the MacArthur-Bates Communicative Development Inventory (Fenson et al., 2007; Finnish version Lyytinen, 2000) at 24 months of age (n = 21), (2) performance at least 1.25 *SD* below age expectations on a screening instrument, the Fox Language Inventory (Korpilahti & Eilomaa, 2002) carried out by a clinical nurse at 36 months (n = 22), (3) performance of -1.25 *SD* or more below age expectations on the Renfrew Word Finding Vocabulary Test at 36 months of age; see Korpilahti, Kaljonen, & Jansson-Verkasalo (2016a, 2016b) for details (n = 7), or (4) speech-language service delivery due to delayed language development according to parent report (n = 17). Due to missing datapoints associated with the cohort study, not all criteria could be used for every child. However, 40 children's late talking status could be confirmed using standardized measures (criteria 1 to 3). Only three children were included solely based on criterion 4, early language service delivery[1]. In total, four children were reported to have a clinical diagnosis of language impairment, according to parent report. A control group with no early risk (typical early development, TED) was recruited from the same cohort. Children invited to the control group exhibited (1) performance between -1 *SD* and +1 *SD* on the MacArthur Bates Communicative Development Inventory at 24 months[2] (2) no known history of late talking or speech-language intervention according to parent report (3) age appropriate performance on standardized tests of language, when scores were available (Korpilahti et al., 2016b).

The recruitment resulted in a total of 79 participants (43 LT, 36 TED) between the ages of 7 and 10 years. All children in both TED and LT groups were required to meet the following criteria: (1) normal hearing based upon a pure tone audiometry screening at 20 dB HL (1 kHz, 2 kHz, and 4 kHz)[3], (2) performance reasoning index (PRI) ≥ 70 as measured by the Finnish version of Wechsler Intelligence Scale for Children (WISC-IV; Wechsler, 2003), (3) Finnish spoken as home language, and (4) no frank emotional, behavioral, motor, intellectual, or neurological disability based on parent report. Two children out of 79 were excluded because of PRI below the criterion. In addition, four participants were excluded due to low accuracy on the experimental task, resulting in a total of 73 participants. Of these 73 participants 38 were identified as late talkers and 35 had typical early language development using the criteria described in the first paragraph.

Because of the varying criteria used to define DLD or SLI in earlier studies, we performed analyses separately using loose (PRI ≥ 70, consistent with DLD definition) and stringent (PRI ≥ 85, consistent with SLI definition) inclusion criteria. The LT and TED groups did not significantly differ in age, $W = 487.5$, $p = .050$ (loose IQ criterion); $W = 372$, $p = .060$ (stringent IQ criterion), socioeconomic status measured by maternal education level, $W = 785.5$,

p = .159 (loose IQ criterion); W = 604, p = .192 (stringent IQ criterion), or IQ measured by PRI, $t(68.84)$ = 1.528, p = .131 (loose IQ criterion); $t(61.16)$ = 1.3294, p = .189 (stringent IQ criterion). All groups are described in Tables 1 and 2.

Late talking status and school aged language abilities

In addition to grouping children into two groups, those with a history of late talking (LT) or typical early development (TED) based on criteria described in the chapter "Participants", language performance was measured at 7–10 years of age. In order to do this, we used Narrative Memory and Comprehension of Instructions subtests of the NEPSY-II (Developmental Neuropsychological Assessment; Korkman, Kirk, & Kemp, 2007), and Vocabulary subtest of the WISC-IV. Based on Finnish norming samples, we calculated Language Index scores that were the means of standard scores on these three subtests. Thus, the Language Index scores were on the scale of mean of 10 and standard deviation of 3. Language skills were treated as a continuous variable (as opposed to grouping children using criteria for DLD and typical language status) to fully represent the variability in abilities. We were interested in a fully representative variability of language outcomes since it has been suggested that even though late talking children achieve age expectations in language abilities by school age, many still remain at a low average level (for review see Rescorla, 2011).

Information about the number of participants in each risk group (LT or TED) performing at and below age expectations as well as demographic information on all participants are presented in Tables 1 and 2. For the sake of clarity and brevity, the tables are formatted using two-level categorical division between age appropriate (TD) versus below age expectations (DLD) performance instead of the full variability in the continuous language abilities measure, which we used in the statistical models used to answer the actual study questions. Language index was lower in LT than TED group, $t(70.97) = 3.023$, $p = .003$ (loose IQ criterion); $t(61.47) = 2.6072$, $p = .011$ (stringent IQ criterion). In LT group, 11 of 38 children (with stringent IQ criterion 7 of 32 children) received Language Index scores lower than 1.25 *SD* below age expectations. Also, in TED group, 3 of 35 children (with stringent IQ criterion 2 of 32 children) had Language Index Scores 1.25 *SD* or more below mean. The majority (71.1 %) of children with LT had normal-range language index scores at school-age. The percentage of TED children with normal-range school-age language skills (91.4 %) was higher than in LT group, but there were 3 also TED children (8.6 %) performing below age expectations at school age. Late bloomers ($M = 9.5$, $SD = 2.3$) performed as a group somewhat worse than their TD peers with no history of late talking ($M = 10.7$, $SD = 2.3$), $t(50) = 1.92$, $p = .061$.

Inhibition task

Stimuli. We used a children's version of the Attention Networks Test (ANT) flanker task (Rueda et al., 2004). The ANT task was originally designed to examine three different aspects of attention: alerting, orienting, and inhibition (also referred to as executive attention) which are thought to be separate dimensions. As our aim was to examine inhibition, we only used the inhibition part of the ANT which is a flanker task. Another advantage of the inhibition part of the ANT is that the number of trials for inhibition component on the task is twice as large as for alerting and orienting yielding to more power in statistical analyses.

The stimuli on children's version of ANT consists of an array of five fish. Inhibition was addressed by presenting the children with congruent and incongruent trials. Half of the trials in the task were congruent, in which all the fish were pointing to the same direction. Half were incongruent, in which the target fish in the middle was pointing to different direction than the other fish (Figure 1). Each participant completed 16 practice trials and 96 trials.

Procedure and experimental design. E-prime software (version 2.0.10.356; Psychology Software Tools, Pittsburgh, PA), laptop screen, and response box were used to present the stimuli and record response time and accuracy. Participants were seated approximately 60 cm from the computer screen and target stimuli were presented in 1.24° (middle fish) and 6.96° (all five fish) visual angle. Participants were instructed to observe the middle fish and press the corresponding button according to which direction the fish was pointing to as soon as possible. During the practice trials, children were given feedback on accuracy and response time to make sure that they understood the instructions. Children were instructed to use index fingers in both hands to press the response box buttons. To measure the effect of congruency in flanker we compared incongruent and congruent trials. Inhibition was reflected by the effect of flanker type, i.e. longer response times for incongruent trials as opposed to congruent trials.

Data analyses

We analyzed response time (RT) performance on the ANT task using R software (R Core Team, 2019) with packages dplyr (Wickham et al., 2019), lme4 (Bates et al., 2015) and lmerTest (Kuznetsova et al., 2017). Packages sjPlot (Lüdecke, 2019), interactions (Long, 2019), and ggplot2 (Wickham, 2016) were used to create tables and figures. The analysis scripts used are included as supplemental material.

Participants with accuracy below 50 % on either congruent or incongruent trials were excluded to ensure all participants complied with the task instructions, resulting in exclusion of 4 participants. To increase statistical power and avoid pitfalls associated with approaches using measures of central tendency (participant mean or median RT), such as means of two populations being the same for different distributions yielding to missing a true effect (Whelan, 2008), we examined the whole RT distributions of each participant. However, trials with RTs more than 2 *SD* longer or shorter than participant's mean RT from analyses and all RTs shorter than 100 ms were excluded, as well as trials with no response. This criterion filtered out 402 of 7008 trials resulting in 6606 trials. Only RTs for correct responses were included in models, reducing final number of trials to 6310.

To address the research question on whether or not inhibition differed as a function of school age language abilities in children with and without late talking status, we modelled RT as a function of language index (continuous measure), history of late talking and flanker type. Language index scores were scaled and centered around sample mean prior to model fitting to avoid issues in model identification due to very large eigenvalues. Given that accuracy rates were at ceiling (M = 95,5 %, SD = 4,5 %), we focused our analyses on RTs. Participant id (random intercept for participant) within each level of trial type (by-participant random slopes) were used as random factors to take individual variations in RTs (for each flanker type) into account. Since RT distributions were positively skewed, we used log-transformation for RTs which made the distributions normal. We prioritized having both individual RT intercepts and slopes in the model over the distribution fitting because task effect (as reflected by individual slopes) was our main focus. Model assumptions were checked using "check_model"-function from R-package performance (Lüdecke et al., 2020). We hypothesized that 1) the flanker effect would be smaller in the TED group than in the LT group and 2) the flanker effect would be associated with school-age Language Index. Both hypotheses 1) and 2) would be supported, if children with no history of late talking and strong school-age language abilities had the smallest flanker effect.

Results

We first modeled RT as a function of language index, late talking and trial flanker type including children who met the loose non-verbal IQ criterion consistent with the DLD definition, PRI > 70. We found a main effect of flanker type (congruent vs. incongruent), .10, 95 % CI [.08, .12], *t* = 8.30*, p* <.001. Back transformed flanker effect size estimate was 71 ms. The model did not reveal interactions between flanker type and language index or late talking, suggesting that inhibition measured by flanker task was not associated with late talking or school-age language skills. However, there was a significant main effect of language index, -.07, 95 % CI [-.13, -.01], *t* = -2.33*, p* = .020, indicating that RTs were shorter across the entire task in children with stronger as opposed to weaker school-age language abilities. Main effect of late talking for general RTs was not significant, -.05, 95 % CI [-.13, .03], *t* = -1.12, *p* = .264, indicating that general RTs were similar in LT and TED groups. Model summary is presented in Table 3.

We then considered the same model, but included only children who met the stringent non-verbal IQ criterion consistent with the SLI definition, PRI > 85. Results with stricter PRI exclusion criterion were very similar to those with more lenient criteria (Figure 2). Main effect of flanker type was significant, .10, 95 % CI [.07, .12], *t* = 7.72*, p* < .001, back-transformed flanker effect size estimate being 72 ms. There were no significant two- or three-way interactions in this model either, providing no evidence for associations between inhibition and school-age language index or late talking. A main effect of language index was present and very similar to the effect observed in the model using looser inclusion criterion, -.07, 95 % CI [-.14, -.01], *t* = -2.21*, p* = .027. The overall RTs were shorter in children with stronger as opposed to weaker school-age language index. Overall RTs were similar in LT and TED groups, -.04, 95 % CI [-.13, .05], *t* = -.94, *p* = .345. Model summary is presented in Table 4.

Post hoc models

Since investigators have recently chosen to model incongruent trials only when relating language skills to inhibition (Larson et al., 2020), we further modeled RTs from incongruent trials only. We modelled RTs as a function of language index and individual participant intercept as a random factor. Consistent with Larson et al (2020), we found a main effect of language index, -.03, 95 % CI [-.07, .01], $t = -1.67$, $p = .095$ (loose criteria); -.04, 95 % CI [-.08, -.00], $t = -1.86$, $p = .063$ (stringent criteria). Main effect of language was only near significant. This is likely due to less statistical power in these analyses as compared to original models because of a smaller number of trials (3101 vs. 6310 in loose criteria models, 2724 vs. 5551 in stringent criteria models). Importantly, similar to our original model, the effect of language index was similar in a model with congruent trials only, -.03, 95 % CI [-.07, .01], $t = -1.64$, $p = .100$ (loose criteria); -.04, 95 % CI [-.08, -.00], $t = -1.75$, $p = .081$ (stringent criteria). Again, the statistical significance in these models were lower as compared to our original models due to less trials and thus less statistical power.

Discussion

In this study, we sought to investigate the relationship between school-aged language index, late talking risk status and efficiency of inhibition. Our hypothesis was that weak language skills would be associated with restrictions in inhibition in both children with and without history of late talking. We assumed that such weakness would be observable as a pronounced effect of congruency in flanker task when children were school aged. According to the dimensional account of late talking we expected to find an interaction between late talking and flanker effect indicating bigger flanker effect on late talkers as compared to TD children.

Flanker effect, reflecting inhibition, was present in our participants suggesting that the flanker task did function as intended, children were in deed distracted by the incongruent flankers. However, we found no interactions between school-aged language abilities and inhibition. Thus, our models did not reveal evidence for difference in the size of flanker effect in LT and TED groups. Nor did the models reveal three-way interactions between late talking, school-age language index, and inhibition. In conclusion, counter to our hypotheses, we did not find differences in inhibition skills between children with and without history of late talking nor children with high and low current language status.

How do we interpret the null results? One possibility is that the null results reflect the time point at which we measured inhibition. We did not have the opportunity to measure inhibition early on when children's late talking was present. Difficulties in inhibition could have been observable during earlier years of development and resolved with time together with language difficulties. However, this is unlikely given that the school-age language abilities in the LT group were not yet comparable with the TED group.

Second possibility is that the null results reflect our sample of late talkers, few of which exhibited persisting language difficulties. Earlier studies with study groups of children with DLD have reported difficulties in inhibition associated with language difficulties (e. g. Henry et al., 2012; Park et al., 2019). Based on these findings, we expected to observe an interaction between late talking, language index and flanker effect so that flanker effect would have been larger in late talkers with persistent language difficulties as compared to LT and TED peers with typical school-age language status. In line with the dimensional account on late talking we also expected that the flanker effect would be smaller in TED group than in LB group. However, relatively few children could be conceived to meet the diagnostic criteria for DLD (14 children of which 11 with the history of late talking) and even fewer had the clinical diagnosis of language disorder (n = 4, all late talkers). Unfortunately, data from only four children with the clinical diagnosis of DLD were likely too small to have sufficient statistical power to detect the possibility of different results with participants with a DLD diagnosis. Had there been more children with DLD, perhaps we had observed that these children in particular would have had deficits in inhibition. It is conceivable that deficits in inhibition might be associated with either better clinical identification or more severe language disorder. Nevertheless, our results are representative of children with and without late talking history (see Rescorla, 2011). Children with the history of late talking as a group had lower school-age language skills than their TED peers especially in tasks involving grammatical skills, yet most children with the history of LT had age appropriate language skills by school age (74 % of children by kindergarten or first grade and 84 % by second grade in Rescorla's study compared to 78 % in our study group). The third explanation of the results pertains to our task choice in measuring inhibition, namely the flanker task. The meta-analysis by Ebert and Kohnert (2011) suggests that inhibition difficulties are associated with DLD. However, even though some studies have reported pronounced flanker effect in DLD (Ebert et al., 2019; Park et al., 2019), others have not found significant group differences between DLD and TD groups in flanker effect (Arbel & Donchin, 2014; Yang & Gray, 2017). The restrictions in statistical power needed to detect small differences are one possible challenge in using flanker task to measure inhibition. Flanker effect size has been estimated to be approximately 75 ms (Fan et al., 2002), in our study the estimate was 71 - 72 ms. We could assume that possible differences in the size of the effect between groups or individuals are smaller than the effect itself. The power needed for statistical models (in practice the number of participants and especially per-subject trials) to detect "statistically significant" differences increases when the size of the effect decreases. Because we could not base our hypothesis about the effect size between group comparisons on flanker effect on previous LT research (to our knowledge there are no studies published), we based our hypothesis on published DLD inhibition studies using similar task that had found group differences.

Because model simulations for mixed effects models require random structure parameters that were not available from previous studies, we used power calculations for a fixed effects linear model to estimate sufficient number of participants for our study. For 75 participants and a medium effect size, these calculations suggested a power of 85.1 % for detecting differences between DLD and typical groups in flanker effect. As we used continuous data for language abilities instead of dichotomous, the power for detecting an effect of this size would be even higher in these comparisons (Altman & Royston, 2006). On the other hand, possible differences between LT and TED groups in the flanker effect, if such differences were to exist in the first place, are possibly smaller than differences between DLD and control groups. Therefore, it is possible that we missed a small difference between the LT and TED groups due to statistical power. Individual variation in RTs is wide (Rouder & Haaf, 2019) and has not been taken into account in previous studies of flanker effect in DLD because these studies have used aggregated measures of RTs (individual means or medians). This may have resulted in over-confidence in false positive results in prior studies that did not consider item level RTs in random effects of the model. Our choice of using a random structure accounting for individual variance in RTs minimizes the risk for type 2 error but might yield to some loss on statistical power (Matuschek et al., 2017). However, we are confident that a between-groups effect large enough to have practical significance would have been detected with our analyses.

General response times and language

Even though our main focus was in examining inhibition and its associations with language measures, the analyses revealed an effect of school age language index on overall RTs. Are longer general response times associated with weaker school-age language abilities still evidence supporting a connection between inhibition and language skills or something else? One possible explanation for longer RTs associated with weaker language skills could be sustained attention. The overall response times may reflect some aspects of attention even though we found no evidence supporting the relationship of the flanker effect and language measures. For instance, it is possible that children with weaker language skills have difficulties in maintaining vigilance across the task. Maintaining accurate responses across trials requires effort from the participant which could result in slowing of processing when the participant sacrifices speed to maintain accuracy. As Ebert & Kohnert (2011) point out, many tasks designed to measure different aspects of cognition anyhow contain a component of sustained attention. However, if the association between processing speed and school-age language index was due to inhibition weakness, we should have observed stronger association in incongruent trials which place more demands on inhibition than congruent trials. This was not the case. The association between the response times and language skills was similar in both trial types. We therefore suggest that it is unlikely that overall longer RTs in the flanker task reflect inhibition.

Similarly, we cannot fully rule out the possibility that long RTs on flanker task in children with language difficulties are affected by slowness of motor planning, which has been suggested to be impaired in DLD (Sanjeevan et al., 2015). However, evidence from comparing different tasks suggests that slowing related to language difficulties is more pronounced on tasks requiring more operations, of which all do not involve motor processing (Kail, 1994; Miller et al., 2001).

Since the measure of general response times is not a difference score like the flanker effect, it is as a measure statistically more powerful which makes it more likely to detect even small differences between participants and groups. Task effect (such as flanker) estimates are often calculated as the difference of mean RTs for different trial types. The measures being compared, in this case RTs for different trial types, often correlate strongly, indicating shared variance of the measures. Shared variance yields to reliability of difference scores being tremendously lower than the reliability of the component measures of trial type themselves (Draheim et al., 2019). Presumably to overcome the problems associated with measuring inhibition using differences between congruent and incongruent conditions, Larson et al. (2020) in their recent study chose to model response times on incongruent flanker trials only and interpreted a significant difference between children with DLD and controls to reflect differences in inhibition. While acknowledging the pitfalls associated with difference scores, this was not our primary modeling strategy. It fails to consider potential individual differences in general processing speed and may mistake individual differences in inhibition with differences in general processing speed. Our modeling that included both congruent and an incongruent trials and main effects of late talking history and school age language index allowed us to consider the possibility that language abilities are not associated with inhibition, but general processing speed instead. While our post hoc analyses revealed that our results were consistent with the Larson et al. (2020) findings – longer RTs in incongruent trials were associated with weaker language skills – our analysis allowed us to consider reaction times across the entire task including congruent trials. Association of RTs and language skills was similar in both congruent and incongruent trials indicating that the association was not specific to trials thought to measure inhibition.

Ultimately, we considered it most likely that the association between individual differences in RTs over all trial types and school-age language outcomes reflect general processing speed. The generalized slowing hypothesis of DLD (Kail, 1994) suggests language learning can be degraded because of slow processing speed across different modalities. Children with DLD exhibit longer RTs than TD peers on a wide range of different tasks (both linguistic and non-linguistic). This finding has been replicated in many DLD studies using different tasks (e. g. Leonard et al., 2007; Miller et al., 2006; Park et al., 2020). Interpreting the findings from this perspective suggests that slow processing speed rather than poor inhibition is associated with unfavorable school age language outcomes. While prior studies using linguistic stimuli have reported predictive value of response times to language outcomes on toddlers (Fernald et al., 2006; Marchman et al., 2019; Newbury et al., 2016; Peter et al., 2019), to our knowledge, ours if the first study reporting a relationship between non-linguistic processing speed and language skills in late talkers.

It is important, however, to note that we found no difference between late talkers and children with typical early development on overall RT. One might have postulated that the dimensional account of late talking would be supported if late talkers as a group were slower than the TED group. This was not the case. The relationship between language index and general response speed did not differ between children with and without the history of late talking, suggesting that factors above and beyond processing speed explain late talking.

Given that school-age language index did not interact with LT status, it is clear that the association between RT and school-age language index is similar in both in children with and without the history of late talking. The association between RT and school-age language skills was observed across the variation of language skills continuum, varying from long RTs in children with weakest language abilities to shortest RTs associated with strongest language abilities including both children with and without the history of late talking.

Finally, motivated by the CATALISE debate on performance IQ criteria (Bishop et al., 2017), we performed data analyses using both lenient and strict IQ inclusion criteria. The results were essentially the same using both criteria, suggesting that the presence or absence of cognitive deficits such as limitations in attention and processing speed are not simply a function of general IQ in children with varying language abilities.

Summary

Contrary to our hypothesis based on previous research on DLD (see Pauls & Archibald, 2016), we found no relationship between inhibition measured by a flanker effect on the ANT tasks and school-aged language index. The flanker effect does not seem to be promising for predicting late talker outcomes.

Despite of not finding associations between inhibition and language index, we found that overall response times on the ANT task were shorter in children with stronger language skills at school age. This effect was similar in children with the history of late talking and children with typical early language development. We therefore suggest that overall response times reflecting general processing speed is a potential predictor for school-age language outcomes in both children with and without late talker history. We are confident that future studies will confirm that early measures of processing speed in late talkers can used to identify children who are likely to exhibit language difficulties that persist into school age.

Acknowledgements

We thank the Steps to the Healthy Development and Well-being of Children Cohort Study for assistance in recruiting children, Pirjo Korpilahti for helpful discussions, Jessica Åkermarck and University of Turku Speech-Language Pathology students (Hanna Fransila, Ella Hölsä, Anna Ikola, Jesmiina Jokinen, Sannamari Kuusisto, Elisa Lavonius, Elina Leminen, Elsa Mannonen, Annina Matikainen, Taina McConchie, Mikaela Mero, Annastiina Miettinen, Susanna Mäenpää, Anni Nuorva, Lotta-Sofia Nurmi, Varpu Oksa, Jenni Paju, Milla Soro, Anniina Sulonen, Outi Sunila, and Kristiina Vahtera) for their assistance in data collection and analysis. Finally, we thank the children and families who participated.

References

Altman, D. G., & Royston, P. (2006). The cost of dichotomising continuous variables. *BMJ*, *332*(7549), 1080.1. https://doi.org/10.1136/bmj.332.7549.1080

Arbel, Y., & Donchin, E. (2014). Error and performance feedback processing by children with Specific Language Impairment—An ERP study. *Biological Psychology*, *99*, 83–91. https://doi.org/10.1016/j.biopsycho.2014.02.012

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

Bishop, D. V. M., Snowling, M. J., Thompson, P. A., Greenhalgh, T., & CATALISE consortium. (2016). CATALISE: A Multinational and Multidisciplinary Delphi Consensus Study. Identifying Language Impairments in Children. *PLOS ONE*, *11*(7), e0158753. https://doi.org/10.1371/journal.pone.0158753

Bishop, Dorothy V.M., Snowling, M. J., Thompson, P. A., Greenhalgh, T., & the CATALISE-2 consortium. (2017). Phase 2 of CATALISE: a multinational and multidisciplinary Delphi consensus study of problems with language development: Terminology. *Journal of Child Psychology and Psychiatry*, *58*(10), 1068–1080. https://doi.org/10.1111/jcpp.12721

Chow, J., Aimola Davies, A. M., Fuentes, L. J., & Plunkett, K. (2019). The vocabulary spurt predicts the emergence of backward semantic inhibition in 18-month-old toddlers. *Developmental Science*, *22*(2). https://doi.org/10.1111/desc.12754

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

Draheim, C., Mashburn, C. A., Martin, J. D., & Engle, R. W. (2019). Reaction time in differential and developmental research: A review and commentary on the problems and alternatives. *Psychological Bulletin*, *145*(5), 508–535. https://doi.org/10.1037/bul0000192

Ebert, K. D., & Kohnert, K. (2011). Sustained Attention in Children With Primary Language Impairment: A Meta-Analysis. *Journal of Speech, Language, and Hearing Research*, *54*(5), 1372–1384. https://doi.org/10.1044/1092-4388(2011/10-0231)

Ebert, K. D., Rak, D., Slawny, C. M., & Fogg, L. (2019). Attention in Bilingual Children With Developmental Language Disorder. *Journal of Speech, Language, and Hearing Research*, *62*(4), 979–992. https://doi.org/10.1044/2018_JSLHR-L-18-0221

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, *16*(1), 143–149. https://doi.org/10.3758/BF03203267

Fan, J., McCandliss, B. D., Sommer, T., Raz, A., & Posner, M. I. (2002). Testing the Efficiency and Independence of Attentional Networks. *Journal of Cognitive Neuroscience*, *14*(3), 340–347. https://doi.org/10.1162/089892902317361886

Feldman, H. M., Dale, P. S., Campbell, T. F., Colborn, D. K., Kurs-Lasky, M., Rockette, H. E., & Paradise, J. L. (2005). Concurrent and Predictive Validity of Parent Reports of Child Language at Ages 2 and 3 Years. *Child Development*, *76*(4), 856–868. https://doi.org/10.1111/j.1467-8624.2005.00882.x

Fenson, L., Marchman, V. A., Thal, D. J., Dale, P. S., Reznick, J. S., & Bates, E. (2007). *MacArthur-Bates Communicative Development Inventories: User's guide and technical manual* (2nd ed.). Paul H. Brookes Publisihing.

Fernald, A., Perfors, A., & Marchman, V. A. (2006). Picking up speed in understanding: Speech processing efficiency and vocabulary growth across the 2nd year. *Developmental Psychology*, *42*(1), 98–116. https://doi.org/10.1037/0012-1649.42.1.98

Finneran, D. A., Francis, A. L., & Leonard, L. B. (2009). Sustained Attention in Children With Specific Language Impairment (SLI). *Journal of Speech, Language, and Hearing Research*, *52*(4), 915–929. https://doi.org/10.1044/1092-4388(2009/07-0053)

Fisher, E. L. (2017). A Systematic Review and Meta-Analysis of Predictors of Expressive-Language Outcomes Among Late Talkers. *Journal of Speech, Language, and Hearing Research*, *60*(10), 2935–2948. https://doi.org/10.1044/2017_JSLHR-L-16-0310

Gathercole, S. E., & Baddeley, A. D. (1990). Phonological memory deficits in language disordered children: Is there a causal connection? *Journal of Memory and Language*, *29*(3), 336–360. https://doi.org/10.1016/0749-596X(90)90004-J

Henry, L. A., Messer, D. J., & Nash, G. (2012). Executive functioning in children with specific language impairment: Executive functioning and SLI. *Journal of Child Psychology and Psychiatry*, *53*(1), 37–45. https://doi.org/10.1111/j.1469-7610.2011.02430.x

Im-Bolter, N., Johnson, J., & Pascual-Leone, J. (2006). Processing Limitations in Children With Specific Language Impairment: The Role of Executive Function. *Child Development*, *77*(6), 1822–1841. https://doi.org/10.1111/j.1467-8624.2006.00976.x

Kail, R. (1994). A Method for Studying the Generalized Slowing Hypothesis in Children With Specific Language Impairment. *Journal of Speech, Language, and Hearing Research*, *37*(2), 418–421. https://doi.org/10.1044/jshr.3702.418

Kaushanskaya, M., Park, J. S., Gangopadhyay, I., Davidson, M. M., & Weismer, S. E. (2017). The Relationship Between Executive Functions and Language Abilities in Children: A

Latent Variables Approach. *Journal of Speech, Language, and Hearing Research*, *60*(4), 912–923. https://doi.org/10.1044/2016_JSLHR-L-15-0310

Korkman, M., Kirk, U., & Kemp, S. (2007). *NEPSY-II: A Developmental Neuropsychological Assessment* (2nd ed.). Harcourt Assessment.

Korpilahti, P., & Eilomaa, P. (2002). *Kettu-testi, Puheen ja kielen arviointimenetelmä [The Fox Language Inventory]* (2nd ed.). Invapaino.

Korpilahti, P., Kaljonen, A., & Jansson-Verkasalo, E. (2016a). Identification of biological and environmental risk factors for language delay: The Let's Talk STEPS study. *Infant Behavior and Development*, *42*, 27–35. https://doi.org/10.1016/j.infbeh.2015.08.008

Korpilahti, P., Kaljonen, A., & Jansson-Verkasalo, E. (2016b). Population-Based Screening for Language Delay: Let's Talk STEPS Study. *Psychology*, *07*, 205–214. https://doi.org/10.4236/psych.2016.72023

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13

Lagström, H., Rautava, P., Kaljonen, A., Räihä, H., Pihlaja, P., Korpilahti, P., Peltola, V., Rautakoski, P., Österbacka, E., Simell, O., & Niemi, P. (2012). Cohort Profile: Steps to the Healthy Development and Well-being of Children the STEPS Study. *International Journal of Epidemology*, *42*, 1273–1284. https://doi.org/10.1093/ije/dys150

Larson, C., Kaplan, D., Kaushanskaya, M., & Weismer, S. E. (2020). Language and Inhibition: Predictive Relationships in Children With Language Impairment Relative to Typically Developing Peers. *Journal of Speech, Language, and Hearing Research*, *63*(4), 1115–1127. https://doi.org/10.1044/2019_JSLHR-19-00210

Lehto, J. E., Juujärvi, P., Kooistra, L., & Pulkkinen, L. (2003). Dimensions of executive functioning: Evidence from children. *British Journal of Developmental Psychology*, *21*(1), 59–80. https://doi.org/10.1348/026151003321164627

Leonard, L. B. (2020). A 200-Year History of the Study of Childhood Language Disorders of Unknown Origin: Changes in Terminology. *Perspectives of the ASHA Special Interest Groups*, *5*(1), 6–11. https://doi.org/10.1044/2019_PERS-SIG1-2019-0007

Leonard, L. B., Ellis Weismer, S., Miller, C. A., Francis, D. J., Tomblin, J. B., & Kail, R. V. (2007). Speed of Processing, Working Memory, and Language Impairment in Children. *Journal of Speech, Language, and Hearing Research*, *50*(2), 408–428. https://doi.org/10.1044/1092-4388(2007/029)

Long, J. A. (2019). *interactions: Comprehensive, User-Friendly Toolkit for Probing Interactions*. https://cran.r-project.org/package=interactions

Lüdecke, D. (2019). *sjPlot: Data Visualization for Statistics in Social Science*. https://doi.org/10.5281/zenodo.1308157

Lüdecke, D., Makowski, D., & Waggoner, P. (2020). *performance: Assessment of Regression Models Performance*. https://CRAN.R-project.org/package=performance

Lyytinen, P. (2000). *Varhaisen kommunikaation ja kielen kehityksen arviointimenetelmä*. Niilo Mäki Instituutti.

Marchman, V. A., Ashland, M. D., Loi, E. C., Adams, K. A., Fernald, A., & Feldman, H. M. (2019). Predictors of early vocabulary growth in children born preterm and full term: A study of processing speed and medical complications. *Child Neuropsychology*, *25*(7), 943–963. https://doi.org/10.1080/09297049.2019.1569608

Marcovitch, S., & Zelazo, P. D. (2009). A hierarchical competing systems model of the emergence and early development of executive function. *Developmental Science*, *12*(1), 1–18. https://doi.org/10.1111/j.1467-7687.2008.00754.x

Marton, K., Kelmenson, L., & Pinkhasova, M. (2007). Inhibition control and working memory capacity in children with SLI. *PSYCHOLOGIA -An International Journal of Psychology in the Orient*, *50*(2), 110–121. https://doi.org/10.2117/psysoc.2007.110

Marton, K., Kovi, Z., & Egri, T. (2018). Is interference control in children with specific language impairment similar to that of children with autistic spectrum disorder? *Research in Developmental Disabilities*, *72*, 179–190. https://doi.org/10.1016/j.ridd.2017.11.007

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing Type I error and power in linear mixed models. *Journal of Memory and Language*, *94*, 305–315. https://doi.org/10.1016/j.jml.2017.01.001

Merriman, W. E. (1999). Competition, attention, and young children's lexical processing. In *The emergence of language.* (pp. 331–358). Lawrence Erlbaum Associates Publishers.

Miller, C. A., Kail, R., Leonard, L. B., & Tomblin, J. B. (2001). Speed of Processing in Children With Specific Language Impairment. *Journal of Speech, Language, and Hearing Research*, *44*(2), 416–433. https://doi.org/10.1044/1092-4388(2001/034)

Miller, C. A., Leonard, L. B., Kail, R. V., Zhang, X., Tomblin, J. B., & Francis, D. J. (2006). Response Time in 14-Year-Olds With Language Impairment. *Journal of Speech, Language, and Hearing Research*, *49*(4), 712–728. https://doi.org/10.1044/1092-4388(2006/052)

Mirsky, A. F., Anthony, B. J., Duncan, C. C., Ahearn, M. B., & Kellam, S. G. (1991). Analysis of the elements of attention: A neuropsychological approach. *Neuropsychology Review*, *2*(2), 109–145.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The Unity and Diversity of Executive Functions and Their Contributions to Complex “Frontal Lobe” Tasks: A Latent Variable Analysis. *Cognitive Psychology*, *41*(1), 49–100. https://doi.org/10.1006/cogp.1999.0734

Montgomery, J., Magimairaj, B., & Finney, M. (2009). Working Memory and Specific Language Impairment: An Update on the Relation and Perspectives on Assessment and Treatment. *American Journal of Speech-Language Pathology / American Speech-Language-Hearing Association*, *19*, 78–94. https://doi.org/10.1044/1058-0360(2009/09-0028)

Moyle, J., Stokes, S. F., & Klee, T. (2011). Early Language Delay and Specific Language Impairment. *Developmental Disabilities Research Reviews*, *17*(2), 160–169. https://doi.org/10.1002/ddrr.1110

Mueller, K. L., & Tomblin, J. B. (2012). Examining the Comorbidity of Language Impairment and Attention-Deficit/Hyperactivity Disorder: *Topics in Language Disorders*, *32*(3), 228–246. https://doi.org/10.1097/TLD.0b013e318262010d

Newbury, J., Klee, T., Stokes, S. F., & Moran, C. (2016). Interrelationships Between Working Memory, Processing Speed, and Language Development in the Age Range 2–4 years. *Journal of Speech, Language, and Hearing Research*, *59*(5), 1146–1158. https://doi.org/10.1044/2016_JSLHR-L-15-0322

Park, J., Miller, C. A., Sanjeevan, T., van Hell, J. G., Weiss, D. J., & Mainela-Arnold, E. (2019). Bilingualism and Attention in Typically Developing Children and Children With

Developmental Language Disorder. *Journal of Speech, Language, and Hearing Research*, 1–14. https://doi.org/10.1044/2019_JSLHR-L-18-0341

Park, J. S., Miller, C. A., Sanjeevan, T., van Hell, J. G., Weiss, D. J., & Mainela-Arnold, E. (2020). Bilingualism and Processing Speed in Typically Developing Children and Children With Developmental Language Disorder. *Journal of Speech, Language, and Hearing Research*, 1–15. https://doi.org/10.1044/2020_JSLHR-19-00403

Pauls, L., & Archibald, L. (2016). Executive Functions in Children With Specific Language Impairment: A Meta-Analysis. *Journal of Speech Language and Hearing Research*, *59*, 1–13. https://doi.org/10.1044/2016_JSLHR-L-15-0174

Peter, M. S., Durrant, S., Jessop, A., Bidgood, A., Pine, J. M., & Rowland, C. F. (2019). Does speed of processing or vocabulary size predict later language growth in toddlers? *Cognitive Psychology*, *115*, 101238. https://doi.org/10.1016/j.cogpsych.2019.101238

R Core Team. (2019). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Rescorla, L. (2009). Age 17 Language and Reading Outcomes in Late-Talking Toddlers: Support for a Dimensional Perspective on Language Delay. *Journal of Speech, Language, and Hearing Research*, *52*(1), 16–30. https://doi.org/10.1044/1092-4388(2008/07-0171)

Rescorla, L. (2011). Late Talkers: Do Good Predictors of Outcome Exist? *Developmental Disabilities Research Reviews*, *17*(2), 141–150. https://doi.org/10.1002/ddrr.1108

Rouder, J. N., & Haaf, J. M. (2019). A psychometrics of individual differences in experimental tasks. *Psychonomic Bulletin & Review*, *26*(2), 452–467. https://doi.org/10.3758/s13423-018-1558-y

Rueda, M. R., Fan, J., McCandliss, B. D., Halparin, J. D., Gruber, D. B., Lercari, L. P., & Posner, M. I. (2004). Development of attentional networks in childhood. *Neuropsychologia*, *42*(8), 1029–1040. https://doi.org/10.1016/j.neuropsychologia.2003.12.012

Sanjeevan, T., Rosenbaum, D. A., Miller, C., van Hell, J. G., Weiss, D. J., & Mainela-Arnold, E. (2015). Motor Issues in Specific Language Impairment: A Window into the Underlying Impairment. *Current Developmental Disorders Reports*, *2*(3), 228–236. https://doi.org/10.1007/s40474-015-0051-9

Simon, J. R., & Wolf, J. D. (1963). CHOICE REACTION TIME AS A FUNCTION OF ANGULAR STIMULUS-RESPONSE CORRESPONDENCE AND AGE. *Ergonomics*, *6*(1), 99–105. https://doi.org/10.1080/00140136308930679

Wechsler, D. (2003). *Wechsler intelligence scale for children* (4th ed.). Pearson, Psychological Corporation.

Westerlund, M., Berglund, E., & Eriksson, M. (2006). Can Severely Language Delayed 3-Year-Olds Be Identified at 18 Months? Evaluation of a Screening Version of the MacArthur–Bates Communicative Development Inventories. *Journal of Speech, Language, and Hearing Research*, *49*(2), 237–247. https://doi.org/10.1044/1092-4388(2006/020)

Whelan, R. (2008). Effective Analysis of Reaction Time Data. *The Psychological Record*, *58*(3), 475–482. https://doi.org/10.1007/BF03395630

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org

Wickham, H., François, R., Henry, L., & Müller, K. (2019). *dplyr: A Grammar of Data Manipulation*. https://CRAN.R-project.org/package=dplyr

Yang, H.-C., & Gray, S. (2017). Executive Function in Preschoolers with Primary Language Impairment. *Journal of Speech, Language, and Hearing Research*, *60*(2), 379–392. https://doi.org/10.1044/2016_JSLHR-L-15-0267

Footnotes

[1] Three children in the LT group whose parents reported early language service receipt did not have a corresponding testing information from the cohort study available from age points 2 or 3 years. Even though this information was missing, we deemed it appropriate to include these children, because they continued to receive language related services at school age and exhibited low language abilities on the school age. The history of delayed language development for these three participants was also confirmed by a speech-language pathologist.

[2] One child in the control group had no MacArthur Bates Communicative Development Inventory data available but performed within normal limits in the Fox Language Inventory, The Renfrew Word Finding Vocabulary Test, and Reynell Developmental Language Scales III language comprehension at the age of 36 months.

[3] Two participants had the hearing level of 30 dB on one ear at 1000 Hz or 4000 Hz in the screening but otherwise passed the screening and had normal hearing according to parent reports.

Table 1

Demographic information and performance on standardized tests for the LT and TED groups when less stringent non-verbal IQ criteria (PRI >= 70) was applied.

	Late talkers			*Typical early development*			*All*
Variable	*Typical n = 27 M (SD)*	*DLD*[1] *n = 11 M (SD)*	*All n = 38 M (SD)*	*Typical n = 32 M (SD)*	*DLD*[1] *n = 3 M (SD)*	*All n = 35 M (SD)*	*n = 73 M (SD)*
Age (months)	*112 (10)*	*107 (11)*	*110 (10)*	*107 (7)*	*109 (10)*	*107 (7)*	*109 (9)*
SES[2]	*1.96 (.71)*	*1.91 (.83)*	*1.95 (.73)*	*2.22 (.83)*	*2.00 (1.00)*	*2.20 (.83)*	*2.07 (.79)*
PRI[3]	*105.3 (16.4)*	*92.9 (13.7)*	*101.7 (16.5)*	*109 (18)*	*93.3 (9.7)*	*107.9 (18.2)*	*104.7 (17.5)*
Language index[4]	*9.5 (2.3)*	*5.2 (.8)*	*8.3 (2.8)*	*10.7 (2.3)*	*5.8 (.4)*	*10.2 (2.6)*	*9.2 (2.8)**
Vocabulary[5]	*10.8 (2.9)*	*4.6 (2.5)*	*9.0 (4.0)*	*10.7 (3.1)*	*7.0 (5.6)*	*10.4 (3.4)*	*9.7 (3.7)*
Comprehension of instructions[6]	*10.2 (2.5)*	*7.5 (3.1)*	*9.4 (2.9)*	*11.9 (2.7)*	*7.0 (3.6)*	*11.5 (3.1)*	*10.4 (3.1)**
Narrative memory[7]	*7.6 (4.6)*	*3.5 (2.4)*	*6.4 (4.5)*	*9.2 (3.9)*	*3.3 (1.5)*	*8.7 (4.1)*	*7.5 (4.4)**

*$p < 0.05$ in group comparison between late talkers and children with typical early development

[1] DLD was defined as language index < -1.25 *SD* from population mean
[2] Maternal education level on scale 1–3
[3] Performance reasoning index, WISC-IV
[4] Mean of the standard scores in three subtests (vocabulary, comprehension of instructions, and narrative memory)
[5] WISC-IV Vocabulary subtest, standard score with *M* = 10 and *SD* = 3
[6] NEPSY-II Comprehension of instructions subtest, standard score with *M* = 10 and *SD* = 3
[7] NEPSY-II Narrative memory subtest, standard score with *M* = 10 and *SD* = 3

Table 2

Demographic information and performance on standardized tests for the LT and TED groups, when more stringent IQ criteria (PRI >= 85) was applied.

	Late talkers			*Typical early development*			*All*
Variable	*Typical* *n* = 25 *M (SD)*	*DLD*[1] *n* = 7 *M (SD)*	*All* *n* = 32 *M (SD)*	*Typical* *n* = 30 *M (SD)*	*DLD*[1] *n* = 2 *M (SD)*	*All* *n* = 32 *M (SD)*	*n* = 64 *M (SD)*
Age (months)	*111 (10)*	*105 (12)*	*110 (10)*	*106 (7)*	*108 (13)*	*106 (7)*	*108 (9)*
SES[2]	*1.88 (.67)*	*2.14 (.90)*	*1.94 (.72)*	*2.23 (.86)*	*1.50 (.71)*	*2.19 (.90)*	*2.06 (.79)*
PRI[3]	*107.4 (15.0)*	*101.6 (7.7)*	*106.2 (13.9)*	*111.5 (15.9)*	*104.5 (9.2)*	*111.1 (15.6)*	*108.6 (14.9)*
Language index[4]	*9.7 (2.3)*	*5.3 (.9)*	*8.7 (2.8)*	*10.7 (2.3)*	*5.7 (.5)*	*10.4 (2.5)*	*9.6 (2.8)**
Vocabulary[5]	*11.0 (3.0)*	*4.7 (3.1)*	*9.6 (3.9)*	*10.7 (3.0)*	*6.5 (7.8)*	*10.4 (3.4)*	*10.0 (3.7)*
Comprehension of instructions[6]	*10.0 (2.5)*	*7.6 (3.9)*	*9.5 (3.0)*	*12.3 (2.1)*	*6.5 (4.9)*	*11.9 (2.6)*	*10.7 (3.0)**
Narrative memory[7]	*8.0 (4.6)*	*3.7 (2.7)*	*7.0 (4.5)*	*9.2 (4.0)*	*4.0 (1.4)*	*8.9 (4.1)*	*8.0 (4.4)*

*$p < 0.05$ in group comparison between late talkers and children with typical early development

[1] DLD was defined as language index < -1.25 *SD* from population mean
[2] Maternal education level on scale 1–3
[3] Performance reasoning index, WISC-IV
[4] Mean of the standard scores in three subtests (vocabulary, comprehension of instructions, and narrative memory)
[5] WISC-IV Vocabulary subtest, standard score with *M* = 10 and *SD* = 3
[6] NEPSY-II Comprehension of instructions subtest, standard score with *M* = 10 and *SD* = 3
[7] NEPSY-II Narrative memory subtest, standard score with *M* = 10 and *SD* = 3

Table 3

Model with the inclusion criteria PRI >= 70

	RT (log transformed ms)				
Predictors	*Estimates*	*95 % CI*	*t-value*	*p*	*Estimate back-transformed to ms (95 % CI)*
(Intercept)	6.52	6.47 – 6.58	215.13	**<0.001**	679 (645 - 721)
FlankerType [incongruent]	0.10	0.08 – 0.12	8.30	**<0.001**	71 (56 - 86)
risk [Late talker]	-0.05	-0.13 – 0.03	-1.12	0.264	-34 (-83 - 20)
Language index (centered)	-0.07	-0.13 – -0.01	-2.33	**0.020**	-46 (-83 - 7)
FlankerType [incongruent] * risk [Late talker]	0.02	-0.01 – 0.05	1.39	0.164	13 (-7 - 34)
FlankerType [incongruent] * Language index (centered)	0.02	-0.00 – 0.04	1.66	0.097	13 (0 - 27)
risk [Late talker] * Language index (centered)	0.06	-0.02 – 0.14	1.43	0.152	41 (-14 - 102)
(FlankerType [incongruent] * risk [Late talker]) * Language index (centered)	-0.03	-0.06 – 0.00	-1.83	0.067	-20 (-40 - 0)
Random Effects					
σ^2	0.05				
$\tau_{00\ id}$	0.03				
$\tau_{11\ id.FlankerTypeincongruent}$	0.00				
$\rho_{01\ id}$	-0.18				
ICC	0.35				
N id	73				
Observations	6310				
Marginal R^2 / Conditional R^2	0.066 / 0.396				

Table 4

Model with inclusion criteria PRI >= 85

	RT (log transformed ms)				
Predictors	*Estimates*	*95 % CI*	*t-value*	*p*	*Estimate back-transformed to ms (95 % CI)*
(Intercept)	6.53	6.46 – 6.59	197.16	**<0.001**	685 (639 - 728)
FlankerType [incongruent]	0.10	0.07 – 0.12	7.72	**<0.001**	72 (50 - 88)
risk [Late talker]	-0.04	-0.13 – 0.05	-0.94	0.345	-26 (-83 - 36)
Language index (centered)	-0.07	-0.14 – -0.01	-2.21	**0.027**	-46 (-89 - -6)
FlankerType [incongruent] * risk [Late talker]	0.02	-0.01 – 0.05	1.14	0.255	14 (-6 - 36)
FlankerType [incongruent] * Language index (centered)	0.02	-0.01 – 0.04	1.29	0.199	14 (-6 - 28)
risk [Late talker] * Language index (centered)	0.06	-0.03 – 0.15	1.28	0.201	43 (-20 - 111)
(FlankerType [incongruent] * risk [Late talker]) * Language index (centered)	-0.03	-0.06 – 0.01	-1.66	0.096	-20 (-40 - 7)
Random Effects					
σ^2	0.05				
$\tau_{00\ id}$	0.03				
$\tau_{11\ id.FlankerTypeincongruent}$	0.00				
$\rho_{01\ id}$	-0.18				
ICC	0.36				
N_{id}	64				
Observations	5551				
Marginal R^2 / Conditional R^2	0.067 / 0.404				

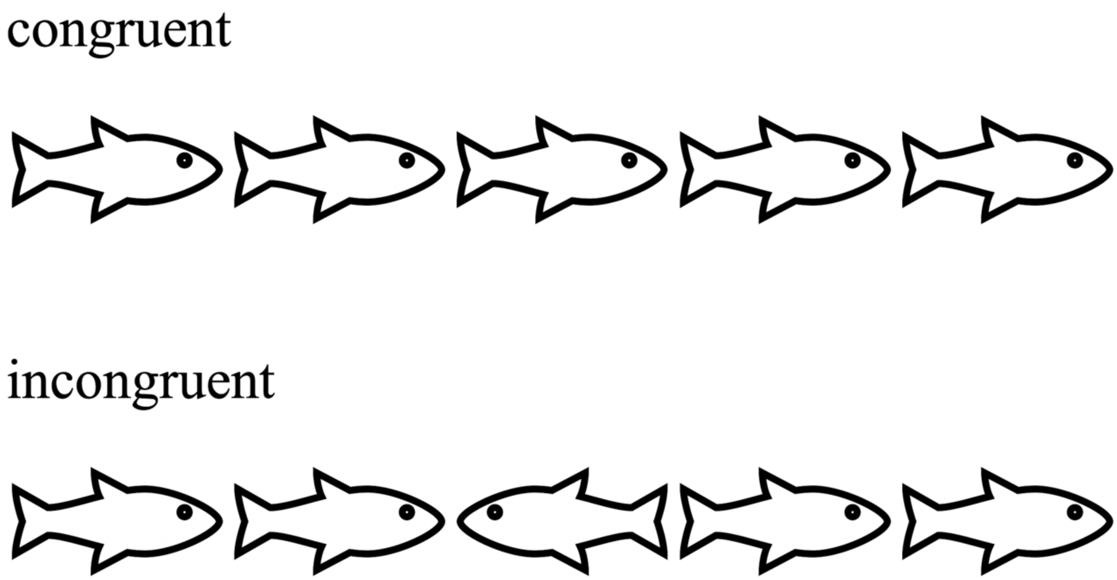

Figure 1: Schematic representation of the congruent and incongruent trials on the Flanker task.

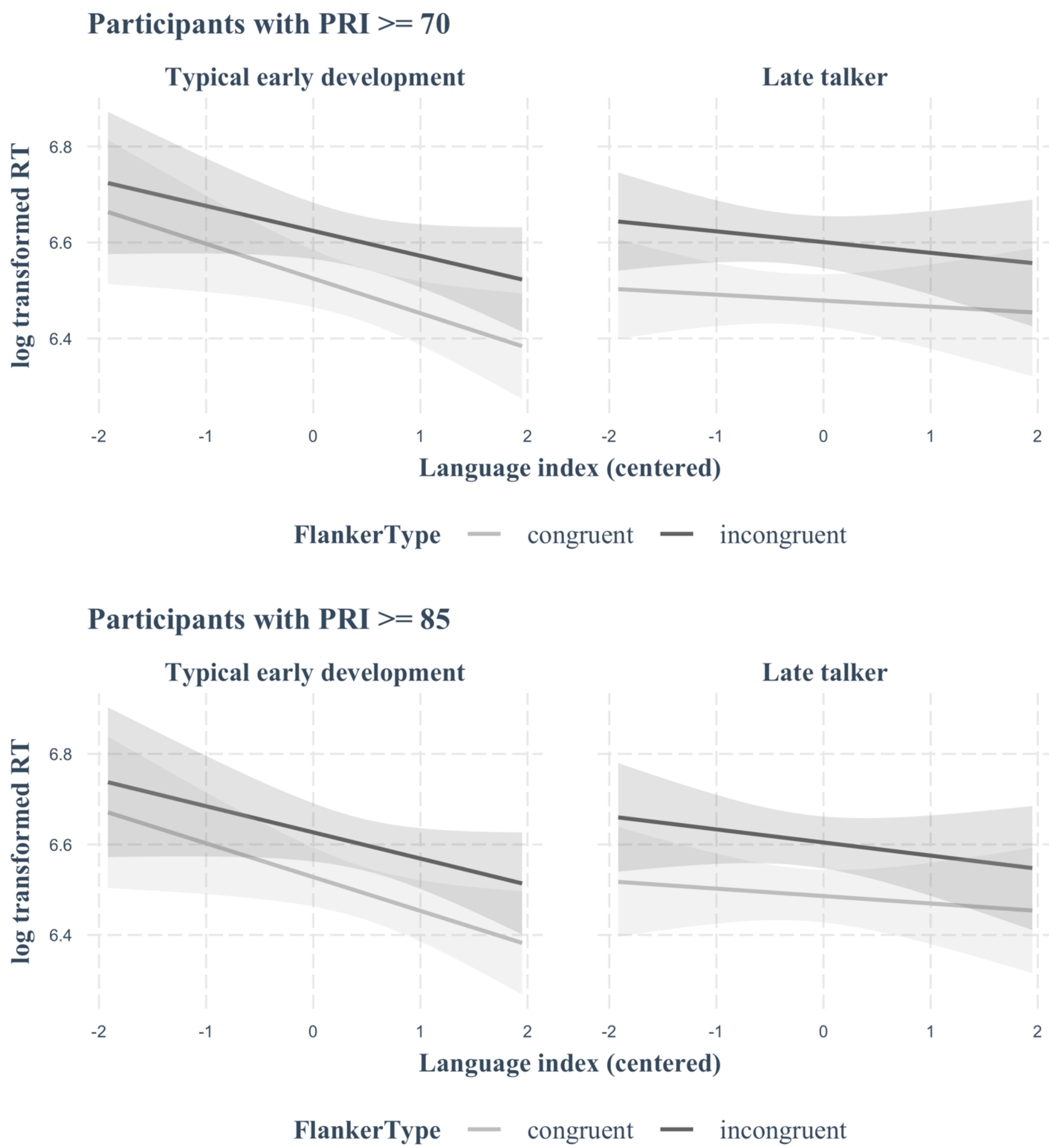

Figure 2: Modeled estimates for response times in the Flanker task with the participant inclusion criteria PRI ⩾ 70 (above) and PRI ⩾ 85 (below).

Supplemental material: https://doi.org/10.23641/asha.14226722