Introducing the Intra-Individual Variability Hypothesis in Explaining Individual Differences in Language Development

Kautto, Anna; Railo, Henry; Mainela-Arnold, Elina

Published in:
Journal of Speech, Language, and Hearing Research

DOI:
10.1044/2024_JSLHR-23-00527

Published: 05/08/2024

Document Version
Final published version

ASHA American Speech-Language-Hearing Association

JSLHR

Research Note

Introducing the Intra-Individual Variability Hypothesis in Explaining Individual Differences in Language Development

Anna Kautto,[a,b] Henry Railo,[a] and Elina Mainela-Arnold[a]

[a] Department of Psychology and Speech-Language Pathology, University of Turku, Finland [b] Department of Speech and Language Pathology, Åbo Akademi University, Turku, Finland

ARTICLE INFO

Article History:
Received August 30, 2023
Revision received February 19, 2024
Accepted April 26, 2024

Editor-in-Chief: Julie A. Washington
Editor: James R. Booth

https://doi.org/10.1044/2024_JSLHR-23-00527

ABSTRACT

Purpose: Response times (RTs) are commonly used in studying language acquisition. However, previous research utilizing RT in the context of language has largely overlooked the intra-individual variability (IIV) of RTs, which could hold significant information about the processes underlying language acquisition.
Method: We explored the association between language abilities and RT variability in visuomotor tasks using two data sets from previously published studies. The participants were 7- to 10-year-old children (n = 77).
Results: Our results suggest that increased variability in RTs is associated with weaker language abilities. Specifically, this within-participant variability in visuomotor RTs, especially the proportion of unusually slow responses, predicted language abilities better than mean RTs, a factor often linked to language skills in past research.
Conclusions: Based on our findings, we introduce the IIV hypothesis in explaining individual differences in language development. According to our hypothesis, inconsistency in the timing of cognitive processes, reflected by increased IIV in RTs, degrades learning different aspects of language, and results in individual differences in language abilities. Future studies should further examine the relationship between IIV and language abilities, and test the extent to which the possible relationship is causal.

Language abilities and the rate of language development vary considerably between individuals. The mechanisms underlying the differences remain largely unknown, although several theories have been put forward to explain the variability (see Kidd et al., 2018). Many of these theories have been presented and investigated in the context of typical and atypical or disordered development, but arguably similar processes explain individual differences across the board.

Response Times and Theories on Individual Differences in Language Development

Response time (RT) is a useful measure in studying language and can be measured from verbal or other motor responses, such as button presses or eye movements. For example, the generalized slowing hypothesis (Kail, 1994) suggests that slowing in cognitive processes results in degraded language learning. Long RTs in individuals with language learning difficulties (e.g., Miller et al., 2001) have been interpreted to support this hypothesis, but the findings have been somewhat inconsistent (see Windsor et al., 2001). In addition to generalized slowing, experimental designs utilizing RT measures have been widely used to study different areas of cognition suggested to support language development, such as attention (Ebert & Kohnert, 2011), or procedural learning (West et al., 2021).

Usually, studies have used individual mean RTs in predicting language abilities. Little attention has been paid to the shape of RT distributions, which are typically right-skewed, well represented by an inverse-gaussian or log-normal distribution (Balota & Yap, 2011). In contrast, individual RT variability and distribution shapes have

Correspondence to Anna Kautto: anna.kautto@utu.fi. **Disclosure:** *The authors have declared that no competing financial or nonfinancial interests existed at the time of publication.*

been studied, for example, in aging and neurogenerative disorders, and are thought to hold potential as a predictor for cognitive impairment (e.g., Costa et al., 2019). A growing field of research also supports the relationship of within-participant variability in RTs and attention deficits (Kofler et al., 2013), and attentional deficits are known to be comorbid with language learning difficulties at least in some samples (Mueller & Tomblin, 2012). However, to our knowledge, there are no published studies on intra-individual variability (IIV) of RTs associated with individual differences in language development and its disorders—a gap we began to address in our analysis.

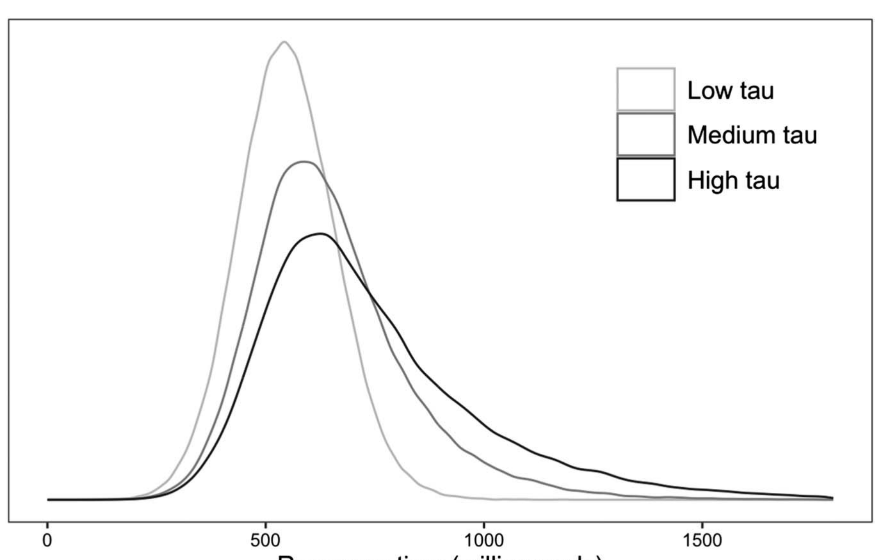
IIV in RTs

IIV refers to within-subject variability in performance in cognitive tasks and can be observed as trial-to-trial fluctuations in RTs. RTs in different modalities exhibit significant IIV—typical visuomotor RT standard deviation in children can reach up to 100 ms, varying according to the task—and this variability also differs substantially between individuals. Variation on this scale may be meaningful in the context of language. It could be used to study the extent to which variability in cognitive processes affect individual's ability to process time-sensitive and rapidly changing signal, such as language.

IIV is often measured by RT standard deviation. Because standard deviation and mean of an RT distribution are typically positively correlated, meaning that longer RTs also tend to have larger variation, a measure of corrected RTs, standard deviation/mean, also called the coefficient of variation in the statistics literature, is sometimes used to account for distribution location (i.e., RT length). However, the use of this measure has been criticized because it may reflect changes in mean RT, variability of RTs, or both (Stawski et al., 2019). A more sophisticated approach, allowing also the characterization of the shape of the distribution, is to fit a known distribution, typically exponentially modified Gaussian (ex-Gaussian), to the RT data. The disadvantage of this method is that individual differences in the shapes of the RT distributions can result in inconclusive parameter values for some individuals. Despite these challenges, it has proven to be a useful measure of individual differences in RTs.

The ex-Gaussian distribution consists of normal and exponential components. It describes RT distributions using three parameters: μ (location), σ (dispersion), and τ (right-skew). The higher the value of τ, the more right-skewed the distribution is (see Figure 1). In the context of studying individual differences, the power of the ex-Gaussian distributional analysis is that especially this τ parameter appears to be relatively stable across different days and has been associated with individual differences in cognitive abilities (Balota & Yap, 2011). Balota and Yap (2011) also mention another useful property of the ex-Gaussian distribution, namely that the mean can be calculated as a sum of μ and τ. Consequently, studies reporting findings related to mean RT may conflate information about location and skewness. The ex-Gaussian approach allows testing if, for example, the observed association between language ability and

Figure 1. Simulated response time distributions with similar values of μ (500) and σ (100) but varying value of τ (low 50, medium 150, high 250).



mean RT is mainly driven by one of these parameters. If the relationship of mean RTs and language indeed reflects generalized slowing, language abilities would be related to ex-Gaussian μ, but if it is driven by the proportion of exceptionally slow responses, τ would be a stronger predictor for language abilities. Ex-Gaussian σ and τ measure different aspects of IIV; the dispersion and skewness of the distribution.

In this research note, we reanalyzed data from two RT studies, in which shorter mean visuomotor RTs were observed to be associated with stronger language abilities (Kautto et al., 2021; Kautto & Mainela-Arnold, 2022). We investigate if the parameters reflecting the location (μ), dispersion (σ), and right-skew (τ) of ex-Gaussian distribution fitted to individual RTs, would be associated to language abilities. Augmenting RT analysis with information about the shape of the distribution could also potentially help resolve the conflicting findings (Windsor et al., 2001) about the relationship of long RTs and language abilities.

Materials and Method

Participants and Language Data

We reanalyzed RT data from two previously published articles reporting findings on Attentional Networks Task (ANT; Kautto et al., 2021) and Serial Response Time task (SRT; Kautto & Mainela-Arnold, 2022). The participants (n = 77, ages 7–10 years) in both studies were the same, recruited from the NeuroTalk study, which is a part of the Southwestern Birth Cohort study (Lagström et al., 2012). All but two participants had data from both tasks. Data for ANT and SRT tasks were collected on separate visits. Language abilities were measured by Language Index (population M = 10, SD = 3, sample M = 8.98, SD = 2.86, range: 3.67–14.67), calculated as a mean of three standardized subtests: Developmental Neuropsychological Assessment (Korkman et al., 2007), Narrative Memory and Comprehension of Instructions subtests, and the Wechsler Intelligence Scale for Children–Fourth Edition (Wechsler, 2003) Vocabulary subtest. Fifteen participants (19.5%) had language abilities 1.25 SD or more below age expectations, thus meeting the diagnostic criteria often used for developmental language disorder, meaning that our sample was well representative of children with weak language abilities. Detailed information on participant recruiting, inclusion and exclusion criteria, and language measures are reported in the previous publications (Kautto et al., 2021; Kautto & Mainela-Arnold, 2022). The cohort study and the NeuroTalk study were both approved by the ethics committee of the Hospital District of Southwest Finland.

RT Tasks

ANT is a combination of cued RT and flanker tasks, designed to investigate three networks of attention: orienting, alerting, and executive attention (Fan et al., 2002). In the children's version of the task the stimuli are five fish, and the participant is asked to press a button corresponding with the orientation of the middle fish. The rest of the fish have either the same orientation (congruent flanker) or opposing orientation (incongruent flanker) as the middle fish. The task also involves cued attention manipulations, measuring the effects of alerting, and orienting. In our ANT task, time between the trials varied from 2,000 to 3,200 ms after the participant's response. In this study, we focus on overall RTs across all study manipulations as the number of trials ($n \leq 24$) for each manipulation type was too low for reliable distribution fitting. However, we calculated individual ex-Gaussian distribution parameters separately for the flanker-congruent and flanker-incongruent trials as the number of trials for these conditions were higher (mean n = 44). Note that in the original study, we did not observe the congruency effect to be related to language abilities (Kautto et al., 2021). The effects of cue type manipulations in relation to language were also not related to language abilities.

SRT was designed to measure pattern learning (Nissen & Bullemer, 1987; Tomblin et al., 2007). In the task, the participants are asked to press a button corresponding to stimulus location. The stimulus in our experiment was a small green creature appearing in one of four horizontally aligned boxes. The task consisted of three blocks of 100 trials, for which the ex-Gaussian distribution was fitted separately. In the two first blocks, the stimulus appearing in the four boxes followed a fixed pattern. In the third block, the stimulus location order was randomized. RT changes during the different blocks of SRT are thought to reflect learning the pattern in the stimulus locations. The participant was asked to respond to the stimuli as soon as possible by pressing a button corresponding to stimulus location. Our task was self-paced, meaning that the stimulus appeared immediately after the participant's response to the previous one.

Distribution Fitting and Statistical Models

R software (R Core Team, 2019) with libraries fitdistrplus (Delignette-Muller & Dutang, 2015), emg (Garbett & Kozdoba, 2020), and actuar (Goulet, 2008) were used in distribution fitting; lme4 (Bates et al., 2015) and party (Hothorn et al., 2006) for statistical modelling; and sjPlot (Lüdecke, 2019) and ggeffects (Lüdecke, 2018) for data visualization. Unfiltered data of correct responses (92% in both tasks) was used for statistical analyses. As the number of trials (> 30) allowed fitting the distribution separately for the three phases of the SRT task and incongruent and congruent trials for ANT, distribution parameters could be calculated to five separate "tasks" for each participant. Four participants' data from ANT incongruent trials were excluded because of very low accuracy (< 50%, also resulting in correct responses $n < 30$).

We calculated parameters for each RT distribution by fitting exponentially modified Gaussian (ex-Gaussian) distribution to the participant's RTs. To address how well the distributions fitted each individual data set, p values for chi-squared test were examined as a measure of the goodness of fit. Of the 377 distribution fits, 96 (25.5%) had p values below .05, indicating bad fit (ANT congruent $n = 20$, ANT incongruent $n = 18$, SRT first pattern phase $n = 18$, second pattern phase $n = 19$, and random phase $n = 21$). As visual inspection of these distributions did not suggest clear differences between the distributions with p values above and below .05, we performed the statistical modeling with all data. This approach was also chosen to avoid excluding certain subgroups of participants whose RT distributions did not follow ex-Gaussian distribution, which could have resulted in an unrepresentative sample. All individual trial counts, distribution fits and chi-squared test p values, as well as the code and results for statistical models, both with and without this exclusion criterion, are available at this website (https://osf.io/jf8b3/).

First, RT standard deviations were calculated for each participant and task condition (ANT flanker type or SRT phase) within participants, resulting in five data points for each participant (two for ANT and three for SRT). We then modelled RT standard deviation with a linear mixed-effect model, language abilities, task type (ANT or SRT), and their interaction as predictors and individual intercept accounting for nested observations as a random factor, lmer notation *RT* standard deviation ~ *Language Index* × *task* + (*1*|*id*). RT standard deviation is methodologically the simplest and also the most used parameter in RT IIV studies, and we wanted to include a simple analysis without the relatively high methodological burden of distribution fitting models. After that, we modelled the parameters reflecting the location (μ), dispersion (σ), and right-skew (τ) of fitted ex-Gaussian distributions separately with linear mixed-effect models. The distribution parameters were used as dependent variables and language abilities as a predictor, with a random intercept to account for within-participant variation, lmer notation *parameter* ~ *Language Index* + (*1*|*id*). The model had to be fitted separately for each parameter because of the high correlation between μ and σ (in our sample, Pearson correlation between μ and σ was .56, $p < .001$ for SRT and .61, $p < .001$ for ANT), and to be able to model data for each subtask while accounting for nested observations for each participant. We also fitted models with trial congruency (flanker incongruent vs. congruent) for ANT and block type (pattern phase vs. random phase) for SRT as predictors but did not observe statistically significant interactions between language ability and congruency/block type. Thus, we report here the simpler models with language abilities as the only predictor. Summaries of the models with interaction terms are available at this website (https://osf.io/jf8b3/).

We used random forest models to compare the predictive power between the distribution parameters. Random forest models are based on a decision tree approach in which an outcome is predicted from data by separating cases in subgroups. The ability to compare the importance between variables despite collinearity has made random forests increasingly popular also in studying language (for applications, see, e.g., Gasparini et al., 2023; Matsuki et al., 2016). We fitted a random forest model utilizing conditional inference trees as base learners (1,500 trees, two variables tried at each split, and a minimum of five observation in each terminal node) to predict language abilities as a function of μ, σ, and τ of RT distributions. The random forest model was fitted separately for the ANT and SRT tasks.

Results

Summary of the RT standard deviation model is presented in Table 1. The model suggested that participants with weaker language abilities exhibited larger standard deviations, especially in the SRT task (see Figure 2). Estimates for changes in the ex-Gaussian parameters for ANT and SRT RTs as a function of Language Index for all six parameter models are summarized in Table 2. For the sake of brevity, full model summaries are only reported in the Open Science Framework repository.

In the ANT task, the effect of σ was significant, weaker language abilities being associated with wider variability of the normal component of RT distributions (see Figure 3, left). In the SRT task, τ was a significant predictor for language abilities (see Figure 3, right), suggesting increased right-skewness to be associated with weaker language abilities. In both ANT and SRT, μ was a near-significant predictor for Language Abilities, weaker language abilities being associated with larger values. We did not perform correction for multiple comparisons in these exploratory models. However, τ in the SRT task would have been a significant predictor for language abilities even with a Bonferroni correction (α-level 3 × .05 = .016).

For the ANT task, the random forest model failed to predict language abilities from the RT distribution parameters, Pearson correlation between out-of-bag

Table 1. Response time standard deviations as a function of language abilities (Language Index) and task (Serial Response Time or Attentional Networks Task [ANT]).

Predictors	Response time *SD*			
	Estimate	**95% CI**	***t***	***p***
(Intercept)	314.63	[267.68, 361.58]	13.18	**< .001**
Language Index	−8.73	[−13.65, −3.80]	−3.48	**.001**
Task (ANT)	−96.08	[−145.84, −46.32]	−3.80	**< .001**
Language Index × Task (ANT)	5.34	[0.13, 10.54]	2.02	**.045**
Random Effects				
σ^2	4955.82			
$\tau_{00\ id}$	2132.03			
ICC	0.30			
N_{id}	77			
Observations	380			
Marginal R^2 / Conditional R^2	0.117 / 0.383			

Note. CI = confidence interval; ICC = intraclass correlation coefficient; boldface = $p < .05$.

predicted and observed values .087, $t(147) = 1.06$, $p = .289$; and comparisons between the predictors were thus not meaningful. For the SRT task, Pearson correlation between out-of-bag predicted and observed values .296, $t(226) = 4.65$, $p < .001$, conditional variable importance suggested τ to clearly be the most informative parameter for predicting language abilities (see Figure 4).

Discussion

In this analysis combining data from two experiments, we observed the relationship between RTs and language index scores to be more clearly associated to variability (standard deviation or ex-Gaussian τ) than length (ex-Gaussian μ) of RTs, suggesting weaker language abilities to

Figure 2. Response time standard deviations as a function of language abilities and task: Serial Response Time (SRT) or Attention Networks Task (ANT).

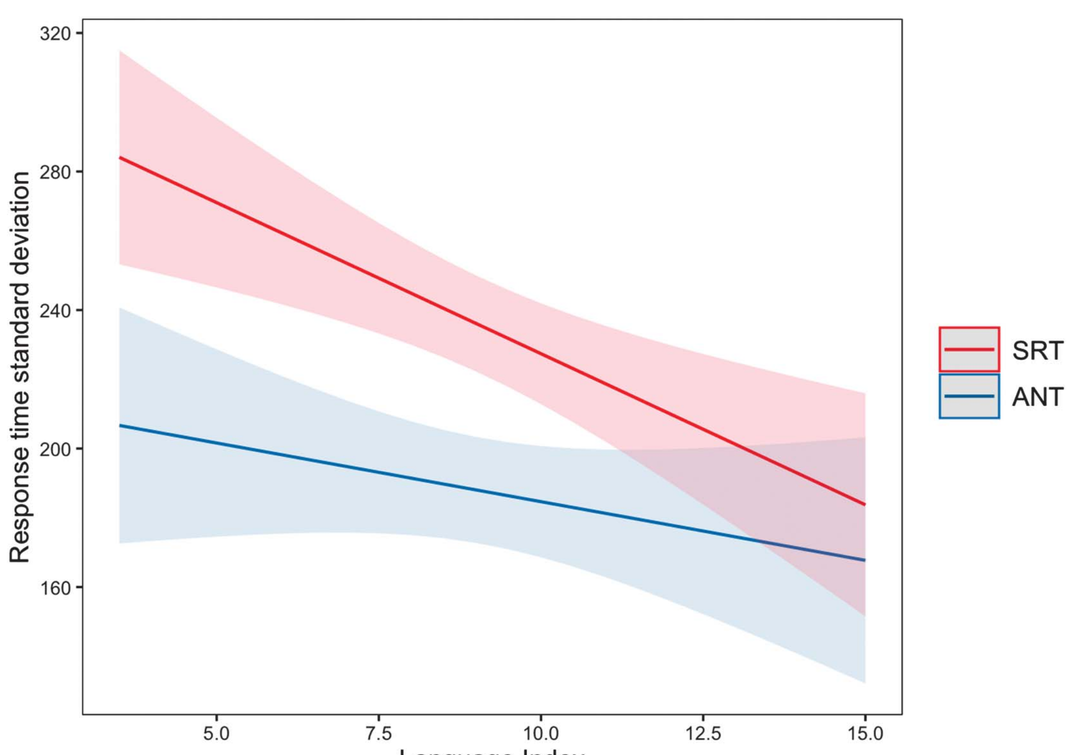


Table 2. Estimates for changes in the ex-Gaussian parameters in the Attentional Networks Task (ANT) and the Serial Response Time (SRT) task as a function of language abilities.

Model	Estimated effect of Language Index			
	Estimate	95% CI	*t*	*p*
μ in ANT	−7.55	[−15.13, 0.03]	−1.97	.051
σ in ANT	−3.21	[−6.28, −0.14]	−2.07	**.041**
τ in ANT	−1.53	[−6.51, 3.44]	−0.61	.544
μ in SRT	−6.77	[−13.57, 0.02]	−1.96	.051
σ in SRT	−1.45	[−3.82, 0.92]	−1.21	.228
τ in SRT	−7.30	[−12.75, −1.84]	−2.64	**.009**

Note. CI = confidence interval; boldfaced = $p < .05$.

be associated with increased IIV. It is noteworthy that the individual RT mean, often used as a measure of RTs, is modulated by the distribution shape. The proportion of exceptionally slow responses (reflected by ex-Gaussian parameter τ) affects the mean so that a participant with a large proportion of these long responses will exhibit a long mean RT, even though their most typical RTs might be much shorter. In previous literature, the mean has often been the only parameter investigated in relation to language abilities, and methodological choices in many studies have resulted in ignoring RT variability. The results of our random forest model suggest that RT variability is even more important factor in predicting language development than the distribution location. Some research findings on response slowness associated with language disorders may be at least partly attributed to the proportion of especially slow responses (ex-Gaussian τ) rather than overall slowness of processing.

In the SRT task, participants with weaker language abilities showed more right-skew in the RT distributions, reflected by ex-Gaussian parameter τ. In the ANT task, participants with weaker language abilities showed more dispersion in the normal component of their RT distributions, reflected by ex-Gaussian parameter σ. Possible explanations for the difference between the tasks are discussed in the Methodological Considerations section.

The IIV Hypothesis of Language Acquisition

Increased IIV in RTs has earlier been observed in relation to attentional difficulties (for meta-analysis, see Kofler et al., 2013), but to our knowledge, this study is

Figure 3. Observed (points) and estimated (lines) values of ex-Gaussian distribution parameters σ (dispersion) in Attention Networks Task (ANT) task (left) and τ (right-skew) for response times in Serial Response Time (SRT) task (right) as a function of Language Index (population $M = 10$, $SD = 3$). The gray area represents 95% confidence intervals.

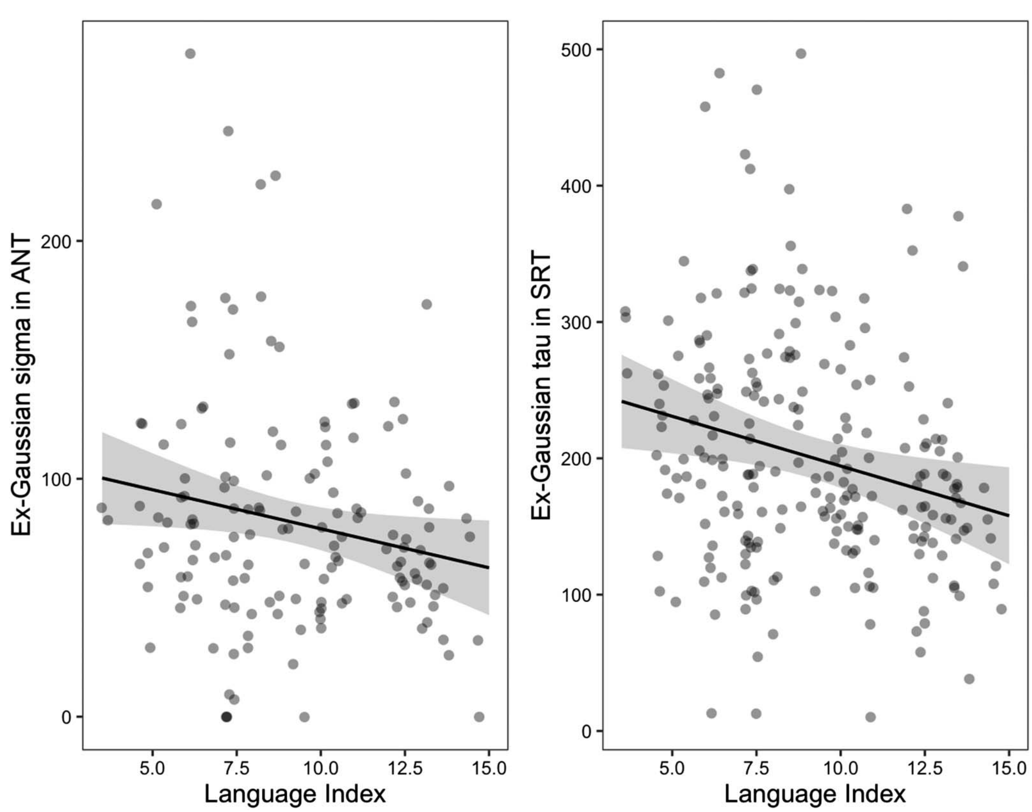
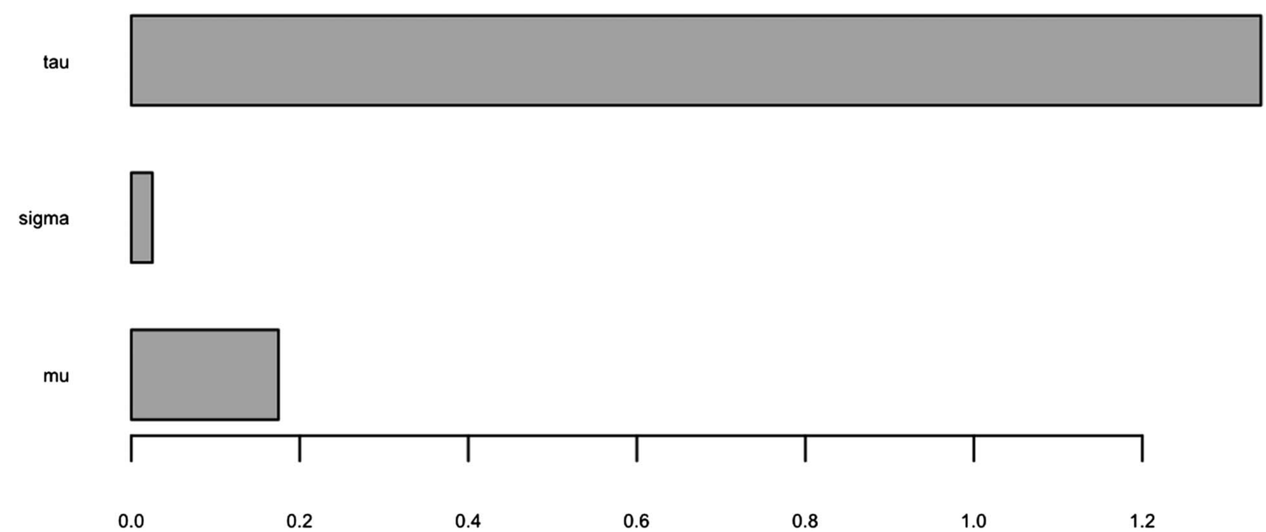


Figure 4. Conditional variable importance (following the permutation principle of the "mean decrease in accuracy") between the three predictors of the random forest model for language abilities: μ (mu), σ (sigma), and τ (tau) from the ex-Gaussian distribution fitted to response time data from the Serial Response Time task.



the first to investigate the relationship between language abilities and IIV in RTs. Based on our findings, we suggest IIV hypothesis of individual differences in language acquisition.

Learning language can be seen as constructing a mental model of language based on sensory input. This input is ideally versatile as language itself, including different sensory modalities and their integration, and conveying information about different aspects of language (such as semantics, morphosyntax, pragmatics, and phonology). However, versatile linguistic environment alone does not guarantee efficient language acquisition if the individual fails to make efficient use of the input and translate it to learning. A central task in learning is updating the model based on input. One example of a phenomenon in which this process can be observed in child language learning is overgeneralizations, for example, a child using the word "car" to refer to all vehicles, or making an inflection error such as "runned" instead of "ran," which are typical in early language development but will disappear later on as a result of further model updating (see Ambridge et al., 2013). If processing the input carrying linguistic meaning is unstable, updating the mental model will take more time and, for example, overgeneralizations in a child's speech may disappear slowly.

Based on the findings of this study, we suggest that large variability in processing linguistic input degrades encoding information and thus language learning, and underlies individual differences in language abilities. This could be compared to listening in noisy environment but the "noise" comes from neurocognitive processes themselves, yielding to low signal-to-noise ratio. The idea is similar to the neural noise hypothesis of dyslexia (Hancock et al., 2017), a condition highly comorbid with developmental language disorder, potentially partly sharing the same mechanisms (Catts et al., 2005). Among cognitive abilities, language is particularly sensitive to unstable input because language is highly time-bound and the signal is very rapidly changing. At the neural level, learning critically depends on simultaneous or near-simultaneous activation of neurons (a modern version of this so called Hebbian learning is known as spike-timing dependent plasticity; Feldman, 2012). This means that variation in the latency of activation of neurons is likely to impair learning.

The IIV hypothesis could help to understand findings supporting the procedural learning hypothesis (Ullman & Pierpont, 2005), which suggests that domain-general difficulties in learning regularities degrade acquiring the rule-governed aspects of language. Evidence from studies of procedural learning suggest that children with weak language abilities differ from their peers especially in their learning rates but learning outcomes might be similar after longer periods of practice (Lum et al., 2014). Slower learning rates in children with language learning difficulties as compared to typically developing peers have been observed, for example, in lexical acquisition (Gray, 2004; Zens et al., 2009). The IIV hypothesis may explain why procedural learning is deficient as both language learning and procedural learning could be affected by the IIV of processing.

Methodological Considerations

In this study, the ANT and SRT data sets had four fundamental differences: (a) The SRT task was self-paced, the next trial appearing immediately after the previous one, while in the ANT task, the interstimulus interval varied between 2,000 and 3,200 ms after the previous response. (b) In the SRT task, stimuli varied only concerning their location (see Kautto & Mainela-Arnold, 2022, for details); whereas in the ANT tasks, there were different trial types: congruent versus incongruent and different cue types (double/center/spatial cues and no cue, see Rueda et al., 2004, for details), and the distributions were fitted for congruent and incongruent trials across cue types. (c) Trial counts for SRT were higher than for ANT. (d) The SRT task had four possible answer buttons whereas the ANT only had two. Any of these factors could potentially affect the fit of the distribution. This also means that there were differences between the trials that were not accounted for in analyzing the data. We ended up fitting the distributions across these potential sources of systematic variance and only fitted the distributions separately for different congruency types. Ideally, we could have fitted distributions separately also for the different cue types, but this would have resulted in maximum 12 trials for each distribution, arguably insufficient for fitting the ex-Gaussian. Despite combining trials with different cue types, the trial counts were significantly smaller for ANT than for SRT. Together, these factors may explain why distribution fits were more accurate for SRT than ANT data and the effects observed in the statistical models were also clearer for SRT. The forementioned differences within ANT trials, together with overall lower number of trials, could have resulted in a difficulty differentiating between the forms of dispersion, reflected by σ and τ. These potential effects should be investigated in future studies.

To some extent, the ex-Gaussian parameters have been suggested to reflect different components of processing; for example, τ has been associated with attentional lapses during a task (Kofler et al., 2013) and individual differences in cognitive abilities, such as working memory (Balota & Yap, 2011). Some studies have suggested the normal component of ex-Gaussian distribution (μ and σ) to reflect decision time and the exponential component (τ) motor RT (Marmolejo-Ramos et al., 2023). However, the ex-Gaussian approach has also been criticized on its loose relationship with cognitive components (Matzke & Wagenmakers, 2009). Our hypothesis on IIV underlying individual differences in language acquisition assumes that this variability can be present at different stages of processing. Future studies should investigate whether individual differences in language acquisition are related to variability in some specific stages of processing, possibly utilizing, for example, Ratcliffe's diffusion decision model (Ratcliff & McKoon, 2008), which has been suggested to differentiate between subcomponents of processing more directly (Matzke & Wagenmakers, 2009).

Many existing data sets allow testing our hypothesis because the analysis of IIV is possible with many experimental designs employing RT measurements, as exemplified by this study reanalyzing data from two separate earlier studies. RT measures also offer potential for linking behavioral and brain-level phenomena. For example, Ribeiro et al. (2016) report an association between spontaneous fluctuations in single trial latencies of visual-evoked potentials (N1) and RTs. Sonuga-Barke and Castellanos (2007) present a hypothesis that specific brain network (default-mode network) interference results in spontaneous attentional fluctuations, which in turn leads to increased RT variability. In relation to attention-deficit/hyperactivity disorder (ADHD), Sonuga-Barke and Castellanos have suggested that non–task-specific default-mode network activity may interfere with goal-directed attention, producing lapses of attention and explaining performance variability in ADHD. As ADHD and language learning difficulties often co-occur, investigating the role of IIV holds promise as a measure for understanding the similarities and differences between the two. Hence, one strength of our hypothesis is that IIV in RTs seems to be relatively straightforwardly linked to neural activity (see also Marmolejo-Ramos et al., 2023) and possibly also to white matter microstructure (McCormick et al., 2023), which could open new perspectives to studying the neural bases of language acquisition.

While we acknowledge the preliminary nature of our findings, our hypothesis is worth careful investigation and holds potential as a predictor for language acquisition and its disorders. In the IIV hypothesis, we suggest that IIV at different stages of information processing underlies individual differences in language acquisition, and that the observed IIV in RTs reflects inconsistency of these processes. In the future, it would be important, for example, to test whether the findings of IIV are limited to visuomotor tasks or more generally observed at different stages of cognitive processes.

Data Availability Statement

Due to ethical restrictions the data are only available from the authors by request.

Acknowledgments

This research was financially supported by the University of Turku Graduate School wages awarded to the first author, an anonymous endowed fund to the University of Turku Speech-Language Pathology, and Kommunalrådet C. G. Sundells Stiftelse funds to the Child language research group in Åbo Akademi University Logopedics. The authors thank the Steps to the Healthy Development and Well-being of Children Cohort Study for assistance in recruiting children and University of Turku Speech-Language Pathology students for their assistance in data collection. Finally, they thank the children and families who participated.

References

Ambridge, B., Pine, J. M., Rowland, C. F., Chang, F., & Bidgood, A. (2013). The retreat from overgeneralization in child language acquisition: Word learning, morphology, and verb argument structure. *Wiley Interdisciplinary Reviews: Cognitive Science, 4*(1), 47–62. https://doi.org/10.1002/wcs.1207

Balota, D. A., & Yap, M. J. (2011). Moving beyond the mean in studies of mental chronometry: The power of response time distributional analyses. *Current Directions in Psychological Science, 20*(3), 160–166. https://doi.org/10.1177/0963721411408885

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Catts, H. W., Adlof, S. M., Hogan, T. P., & Weismer, S. E. (2005). Are specific language impairment and dyslexia distinct disorders? *Journal of Speech, Language, and Hearing Research, 48*(6), 1378–1396. https://doi.org/10.1044/1092-4388(2005/096)

Costa, A. S., Dogan, I., Schulz, J. B., & Reetz, K. (2019). Going beyond the mean: Intraindividual variability of cognitive performance in prodromal and early neurodegenerative disorders. *The Clinical Neuropsychologist, 33*(2), 369–389. https://doi.org/10.1080/13854046.2018.1533587

Delignette-Muller, M. L., & Dutang, C. (2015). fitdistrplus: An R package for fitting distributions. *Journal of Statistical Software, 64*(4), 1–34. https://doi.org/10.18637/jss.v064.i04

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)

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, D. E. (2012). The spike-timing dependence of plasticity. *Neuron, 75*(4), 556–571. https://doi.org/10.1016/j.neuron.2012.08.001

Garbett, S., & Kozdoba, M. (2020). *EMG: Exponentially modified Gaussian (EMG) distribution*. R Foundation for Statistical Computing. https://CRAN.R-project.org/package=emg

Gasparini, L., Shepherd, D. A., Bavin, E. L., Eadie, P., Reilly, S., Morgan, A. T., & Wake, M. (2023). Using machine-learning methods to identify early-life predictors of 11-year language outcome. *The Journal of Child Psychology and Psychiatry, 64*(8), 1242–1252. https://doi.org/10.1111/jcpp.13733

Goulet, V. (2008). actuar: An R package for actuarial science. *Journal of Statistical Software, 25*(7), 1–37. https://doi.org/10.18637/jss.v025.i07

Gray, S. (2004). Word learning by preschoolers with specific language impairment: Predictors and poor learners. *Journal of Speech, Language, and Hearing Research, 47*(5), 1117–1132. https://doi.org/10.1044/1092-4388(2004/083)

Hancock, R., Pugh, K. R., & Hoeft, F. (2017). Neural noise hypothesis of developmental dyslexia. *Trends in Cognitive Sciences, 21*(6), 434–448. https://doi.org/10.1016/j.tics.2017.03.008

Hothorn, T., Hornik, K., & Zeileis, A. (2006). Unbiased recursive partitioning: A conditional inference framework. *Journal of Computational and Graphical Statistics, 15*(3), 651–674. https://doi.org/10.1198/106186006X133933

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

Kautto, A., Jansson-Verkasalo, E., & Mainela-Arnold, E. (2021). Generalized slowing rather than inhibition is associated with language outcomes in both late talkers and children with typical early development. *Journal of Speech, Language, and Hearing Research, 64*(4), 1222–1234. https://doi.org/10.1044/2020_JSLHR-20-00523

Kautto, A., & Mainela-Arnold, E. (2022). Procedural learning and school-age language outcomes in children with and without a history of late talking. *International Journal of Language & Communication Disorders, 57*(6), 1255–1268. https://doi.org/10.1111/1460-6984.12751

Kidd, E., Donnelly, S., & Christiansen, M. H. (2018). Individual differences in language acquisition and processing. *Trends in Cognitive Sciences, 22*(2), 154–169. https://doi.org/10.1016/j.tics.2017.11.006

Kofler, M. J., Rapport, M. D., Sarver, D. E., Raiker, J. S., Orban, S. A., Friedman, L. M., & Kolomeyer, E. G. (2013). Reaction time variability in ADHD: A meta-analytic review of 319 studies. *Clinical Psychology Review, 33*(6), 795–811. https://doi.org/10.1016/j.cpr.2013.06.001

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

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*(5), 1273–1284. https://doi.org/10.1093/ije/dys150

Lüdecke, D. (2018). ggeffects: Tidy data frames of marginal effects from regression models. *Journal of Open Source Software, 3*(26), Article 772. https://doi.org/10.21105/joss.00772

Lüdecke, D. (2019). *sjPlot: Data visualization for statistics in social science*. https://doi.org/10.5281/zenodo.1308157

Lum, J. A. G., Conti-Ramsden, G., Morgan, A. T., & Ullman, M. T. (2014). Procedural learning deficits in specific language impairment (SLI): A meta-analysis of serial reaction time task performance. *Cortex, 51*(100), 1–10. https://doi.org/10.1016/j.cortex.2013.10.011

Marmolejo-Ramos, F., Barrera-Causil, C., Kuang, S., Fazlali, Z., Wegener, D., Kneib, T., De Bastiani, F., & Martinez-Flórez, G. (2023). Generalised exponential–Gaussian distribution: A method for neural reaction time analysis. *Cognitive Neurodynamics, 17*(1), 221–237. https://doi.org/10.1007/s11571-022-09813-2

Matsuki, K., Kuperman, V., & Van Dyke, J. A. (2016). The random forests statistical technique: An examination of its value for the study of reading. *Scientific Studies of Reading, 20*(1), 20–33. https://doi.org/10.1080/10888438.2015.1107073

Matzke, D., & Wagenmakers, E.-J. (2009). Psychological interpretation of the ex-Gaussian and shifted Wald parameters: A diffusion model analysis. *Psychonomic Bulletin & Review, 16*(5), 798–817. https://doi.org/10.3758/PBR.16.5.798

McCormick, E. M., & Kievit, R. A., Cambridge Centre for Ageing and Neuroscience. (2023). Poorer white matter microstructure predicts slower and more variable reaction time performance: Evidence for a neural noise hypothesis in a large lifespan cohort. *The Journal of Neuroscience, 43*(19), 3557–3566. https://doi.org/10.1523/JNEUROSCI.1042-22.2023

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)

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

Nissen, M. J., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology, 19*(1), 1–32. https://doi.org/10.1016/0010-0285(87)90002-8

R Core Team. (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation, 20*(4), 873–922. https://doi.org/10.1162/neco.2008.12-06-420

Ribeiro, M. J., Paiva, J. S., & Castelo-Branco, M. (2016). Spontaneous fluctuations in sensory processing predict within-subject reaction time variability. *Frontiers in Human Neuroscience, 10,* Article 200. https://doi.org/10.3389/fnhum.2016.00200

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

Sonuga-Barke, E. J. S., & Castellanos, F. X. (2007). Spontaneous attentional fluctuations in impaired states and pathological conditions: A neurobiological hypothesis. *Neuroscience & Biobehavioral Reviews, 31*(7), 977–986. https://doi.org/10.1016/j.neubiorev.2007.02.005

Stawski, R. S., MacDonald, S. W. S., Brewster, P. W. H., Munoz, E., Cerino, E. S., & Halliday, D. W. R. (2019). A comprehensive comparison of quantifications of Intraindividual variability in response times: A measurement burst approach. *The Journals of Gerontology: Series B, 74*(3), 397–408. https://doi.org/10.1093/geronb/gbx115

Tomblin, J. B., Mainela-Arnold, E., & Zhang, X. (2007). Procedural learning in adolescents with and without specific language impairment. *Language Learning and Development, 3*(4), 269–293. https://doi.org/10.1080/15475440701377477

Ullman, M. T., & Pierpont, E. I. (2005). Specific language impairment is not specific to language: The procedural deficit hypothesis. *Cortex, 41*(3), 399–433. https://doi.org/10.1016/S0010-9452(08)70276-4

Wechsler, D. (2003). *Wechsler Intelligence Scale for Children* (4th ed.). The Psychological Corporation.

West, G., Melby-Lervag, M., & Hulme, C. (2021). Is a procedural learning deficit a causal risk factor for developmental language disorder or dyslexia? A meta-analytic review. *Developmental Psychology, 57*(5), 749–770. https://doi.org/10.1037/dev0001172

Windsor, J., Milbrath, R. L., Carney, E. J., & Rakowski, S. E. (2001). General slowing in language impairment: Methodological considerations in testing the hypothesis. *Journal of Speech, Language, and Hearing Research, 44*(2), 446–461. https://doi.org/10.1044/1092-4388(2001/036)

Zens, N. K., Gillon, G. T., & Moran, C. (2009). Effects of phonological awareness and semantic intervention on word-learning in children with SLI. *International Journal of Speech-Language Pathology, 11*(6), 509–524. https://doi.org/10.3109/17549500902926881