Research Note

Processing Speed Measures as Clinical Markers for Children With Language Impairment

Jisook Park,^a Carol A. Miller,^a and Elina Mainela-Arnold^b

Purpose: This study investigated the relative utility of linguistic and nonlinguistic processing speed tasks as predictors of language impairment (LI) in children across 2 time points.

Method: Linguistic and nonlinguistic reaction time data, obtained from 131 children (89 children with typical development [TD] and 42 children with LI; 74 boys and 57 girls) were analyzed in the 3rd and 8th grades. Receiver operating characteristic curve analyses and likelihood ratios were used to compare the diagnostic usefulness of each task. A binary logistic regression was used to test whether combined measures enhanced diagnostic accuracy.

Results: In 3rd grade, a linguistic task, grammaticality judgment, provided the best discrimination between LI and TD groups. In 8th grade, a combination of linguistic and nonlinguistic tasks, rhyme judgment and simple response time, provided the best discrimination between groups.

Conclusions: Processing speed tasks were moderately predictive of LI status at both time points. Better LR+ than LR- values suggested that slow processing speed was more predictive of the presence than the absence of LI. A nonlinguistic processing measure contributed to the prediction of LI only at 8th grade, consistent with the view that nonlinguistic and linguistic processing speeds follow different developmental trajectories.

This study investigated how linguistic and nonlinguistic processing speed tasks are related to the diagnosis of language impairment (LI) in children in middle childhood and early adolescence. The term *specific language impairment* (SLI) refers to a developmental language disorder in the absence of any obvious neurological, intellectual, and sensory motor impairment (Leonard, 2014). Individuals with SLI show difficulty in various linguistic areas, including deficits at lexical, morphosyntactic, syntactic, phonological, and pragmatic levels (Leonard, 2014). These difficulties in language development have negative consequences in various areas: academic, emotional, behavioral, and social adjustment; mental health; and vocational outcomes (see Whitehouse, Watt, Line, & Bishop, 2009, for a review). The prevalence rate of SLI is high, approximately 7% for kindergarten students (Tomblin et al., 1997), and SLI is found to persist over time (e.g., Clegg, Hollis, Mawhood, & Rutter, 2005; Johnson et al., 1999;

Tomblin, 2010); however, it often remains undiagnosed and untreated during the school years (Ehren, 2002; Zhang & Tomblin, 2000).

Processing speed refers to the ability to process input before the input decays or before incoming information interferes with it (L. T. Miller & Vernon, 1996), and it has been found to be associated with language abilities (Leonard et al., 2007). Group differences in processing speed between children with SLI and typically developing peers are often small but consistently found across studies and across simple and complex tasks and linguistic and nonlinguistic tasks (Kohnert & Windsor, 2004; C. Miller, Kail, Leonard, & Tomblin, 2001; C. Miller et al., 2006; Windsor, Kohnert, Loxtercamp, & Kan, 2008). However, statistically significant group differences are not always found (e.g., C. Miller et al., 2001). Also, there is mixed evidence on correlations between processing speed and language abilities. For instance, Leonard et al. (2007) found that nonlinguistic and linguistic processing speed predicted language abilities in 14-year-olds, whereas Lahey, Edwards, and Munson (2001) did not find a linear correlation between response time on lexical processing tasks and severity of LI in kindergarten children. Given this variability in the findings, we chose to use a different approach to examine whether slow processing speed on linguistic tasks, nonlinguistic tasks, or both is

^aThe Pennsylvania State University, State College

^bUniversity of Toronto, Ontario, Canada

Correspondence to Jisook Park: jzp151@psu.edu

Editor: Rhea Paul

Associate Editor: Lizbeth Finestack

Received April 2, 2014

Revision received October 11, 2014

Accepted February 4, 2015

DOI: 10.1044/2015_JSLHR-L-14-0092

Disclosure: The authors have declared that no competing interests existed at the time of publication.

characteristic of children with SLI. We used statistical tools for the appraisal of diagnostic accuracy—that is, tools which estimate the degree to which a variable, such as processing speed, is predictive of membership in an affected group.

Methods for appraising diagnostic accuracy include likelihood ratios (Dollaghan, 2007) and logistic regression analyses (Tabachnick & Fidell, 2013). *Positive likelihood ratio* (LR+) refers to how the probability of being affected changes when the test result is positive. Therefore, a high LR+ value is associated with a greater chance that a positive test result identifies true disorder. *Negative likelihood ratio* (LR–) refers to how the probability of being affected changes when the test result is negative. Thus, an LR– value that is close to 0 means a small chance that the negative test result is associated with true disorder. Values of LR+ at or above 10 allow us to diagnose with confidence that an individual has the disorder, whereas LR– values at or below 0.10 enable us to confidently rule out the possibility that an individual has the disorder. Values of LR+ between 3 and 10 and LR– values between 0.3 and 0.1 are considered suggestive (Dollaghan, 2007). The receiver operating characteristic (ROC) curve is recommended to determine the best cutoff score to calculate likelihood ratios. Unlike an arbitrary cutoff, the empirical cutoff provided by the ROC curve enhances the precision of determining diagnostic accuracy for a given clinical marker (Sackett, Haynes, Guyatt, & Tugwell, 1991).

To assess diagnostic usefulness, an index measure is compared with a reference or a gold standard, the best available diagnostic method, to determine if the index measure correctly assigns an individual to a diagnostic category generated by the standard. The EpiSLI system is currently the most systematic, objective, and data-based reference standard available for SLI (Tomblin, Records, & Zhang, 1996). The EpiSLI system used a combination of norm- and criterion-referenced language measures to diagnose children as having SLI. The language criteria for SLI in the EpiSLI system included two or more language scores lower than $SD = 1.25$ below average out of five composite scores: three linguistic areas including vocabulary, grammar, and narrative and two modalities, receptive and expressive. Conversely, the criteria for TD required four or more composite area language scores no more than $SD = 1.25$ below average.

A given processing speed measure might have high diagnostic accuracy at one age point, but poorer diagnostic accuracy at other age points due to developmental changes in processing speed in children. Some researchers have reported discrepancies in developmental trajectories between linguistic and nonlinguistic domains (reviewed in Kail & Miller, 2006). Kail and Miller (2006), using data drawn from the same larger data set as the current study, found that nonlinguistic and linguistic domains had different developmental trajectories in TD children. The nonlinguistic domain showed a steeper developmental slope between 9- and 14-year-olds (third and eighth grade) because children were faster on linguistic tasks than nonlinguistic tasks at age 9 years, but the two domains were similar at age 14 years.

Therefore, we investigated whether nonlinguistic and linguistic processing speed tasks could identify children with LI at these two time points.

We conducted a preliminary study, the first of its kind, using diagnostic accuracy analyses to investigate whether nonlinguistic processing speed tasks in addition to linguistic processing speed tasks could be informative for the prediction of LI status and whether the relationship between nonlinguistic and linguistic processing measures would differ over time in predicting LI. This study takes advantage of an archival data set in which a relatively large population-based sample of children with and without LI were assessed using the EpiSLI system (providing a strong reference standard) and participated in a set of reaction time (RT) tasks, on the basis of previous literature, that varied in the type of nonlinguistic stimulus and in the complexity of the judgment required by the child (C. Miller et al., 2001). In addition, this existing data set provides us with an attractive property for the purposes of the current study because the data set includes processing speed outcomes at two time points (Grades 3 and 8) to consider developmental changes in a longitudinal design.

Method

Participants

The participants were 131 monolingual English-speaking children (74 boys and 57 girls) who received diagnostic testing in the second and eighth grades and participated in RT tasks in the third and eighth grades. These children were a subset of children who participated in a longitudinal study (see Tomblin et al., 1997, 2000, for an overview). The current study included 89 TD children and 42 children with LI who completed RT tasks in both grades and remained in the same diagnostic category from third (8- to 9-year-olds) to eighth grade (13- to 14-year-olds). Previous analyses of the data were originally reported by C. Miller et al. (2001, 2006).

Children who met the following criteria were excluded: (a) those from a second-language environment; (b) those who had a history of other developmental disorders such as autism, intellectual disability, and/or neurological impairments; and (c) those who were blind, used hearing devices, and/or had persistent bilateral hearing deficits.

Nonverbal intelligence was obtained using the Block Design and Picture Completion subtests of the Wechsler Intelligence Scale for Children—Third Edition (WISC-III; Wechsler, 1991). All participants were required to have nonverbal IQs of 75 or higher in both Grades 2 and 8. This cutoff includes some children with nonspecific language impairment (NLI) as defined in C. Miller et al. (2001, 2006). Individuals with NLI and SLI do not qualitatively differ in their language profiles and on linguistic and nonlinguistic processing tasks (Leonard, 2007; C. Miller et al., 2006). They were combined and labeled as LI suggested by Reilly et al. (2014). However, to reduce the probability of including a child with an intellectual disability (conventionally

indicated by $IQ < 70$), a cutoff score of 75 on the nonverbal IQ tests was applied considering the standard error of measurement is ± 5 points for this measure.

Diagnostic language status was based on the EpiSLI system (Tomblin, Records, & Zhang, 1996). The EpiSLI system included three linguistic areas (vocabulary, grammar, and narrative) and two modalities (receptive and expressive). When the children received full diagnostic testing at the second grade, these five composite language measures were derived from several standardized tests and a narrative production task (Fey, Catts, & Proctor-Williams, 2000), as described in Table 1. In eighth grade, the five composite language measures were derived from some of the same tests and others more appropriate for the age group (see Table 1). The children were categorized as having LI if at least two out of the five composite scores were lower than $SD = 1.25$ from the mean and TD if they had at least four out of five composite scores higher than $SD = 1.25$ below the mean (see Table 1). There were 81 White, five African American, one Hispanic, and two Asian children in the TD group, whereas there were 33 White and nine African American children in the LI group.

Processing Speed Tasks

A laptop computer running custom software was used to present the RT tasks. Reaction times were recorded

from a stimulus onset to a voice key for the picture-naming task or to a key press on the keyboard for all the other tasks. Before each task, children completed practice items. More details about the stimuli and procedure can be found in C. Miller et al. (2001, 2006).

Nonlinguistic tasks. Children completed five nonlinguistic speed tasks.

1. The tapping task (TP) required children to tap a key as rapidly as possible with the preferred hand in 5 s.
2. The simple response time task (SR) involved the children striking a key marked with a colored dot as quickly as they could when a visual signal was displayed. The visual signal was an image of three asterisks that appeared with a random delay variation of 1, 2, or 5 s after the written word "ready."
3. In the visual search task (VS), the children were presented with a target figure on one side of the screen and scanned five figures from left to right on the other side of the screen. They pressed one key when the target was present among the five figures and another key when the target was not present.
4. To complete the mental rotation task (MR), children pressed one key when a target figure on the left was the same as a figure on the right. The figure on the

Table 1. Group characteristics in the diagnostic phase.

Grades	Measures	TD (n = 89)	LI (n = 42)	t
2	Nonverbal IQ SS ^a	104.07 (10.58)	90.88 (10.09)	-6.75**
	Receptive composite z ^b	-0.08 (0.76)	-1.73 (0.63)	-12.28**
	Expressive composite z ^c	-0.14 (0.81)	-1.60 (0.47)	-10.79**
	Vocabulary composite z ^d	-0.10 (0.86)	-1.56 (0.71)	-9.54**
	Sentence composite z ^e	-0.19 (0.74)	-1.68 (0.46)	-11.92**
	Narrative composite z ^f	0.02 (0.92)	-1.43 (0.61)	-9.29**
	Years of mother's education	14.07	13.10	-2.83*
	Sex (male/female)	48/41	26/16	
	Race/ethnicity (White/African American/Asian/Hispanic)	81/5/2/1	33/9/0/0	
8	Nonverbal IQ SS ^a	102.71 (9.76)	94.05 (12.74)	-4.28*
	Receptive composite z ^h	0.02 (0.74)	-1.55 (0.59)	-11.99*
	Expressive composite z ⁱ	-0.12 (0.81)	-1.55 (0.60)	-10.18*
	Vocabulary composite z ^j	-0.06 (0.82)	-1.28 (0.48)	-8.91*
	Sentence composite z ^k	-0.02 (0.65)	-1.85 (0.79)	-14.06*
	Narrative composite z ^l	-0.06 (0.97)	-0.99 (0.48)	-5.93*

^aNational standard score of the performance scale of the Wechsler Intelligence Scale for Children-Third Edition (WISC-III). ^bComposite z score of Peabody Picture Vocabulary Test-Revised (PPVT-R) and subtests of Clinical Evaluation of Language Fundamentals-Third Edition (CELF-3): Sentence Structures, Concepts and Directions, Word Structure, Listening to Paragraphs. ^cComposite z score of the Expressive Vocabulary subtest of Comprehensive Receptive and Expressive Vocabulary Test (CREVT); Recalling Sentences, a subtest of CELF-3; and a narrative production task (Fey, Catts, & Proctor-Williams, 2001). ^dComposite z score of PPVT-R and the Expressive Vocabulary subtest of CREVT.

^eComposite z score of subtests of CELF-3: the Sentence Structure, Concepts and Directions, Word Structure, and Recalling Sentences.

^fComposite z score of the Listening to Paragraphs, a subtest of CELF-3 and a narrative production task (Fey, Catts, & Proctor-Williams, 2000).

^gNational standard score of the performance scale of Block Design and Picture Completion, subtests of the WISC-III. ^hComposite z score of PPVT-R; Concepts and Following Directions, a subtest of CELF-3; a narrative comprehension task, a subtest of the Qualitative Reading Inventory-3 (QRI-3).

ⁱComposite z score of the Expressive Vocabulary, a subtest of CREVT; Recalling Sentences, a subtest of CELF-3; and a narrative production task, a subset of the QRI-3. ^jVocabulary composite z = composite z score of PPVT-R and the Expressive Vocabulary, a subtest of CREVT. ^kComposite z score of Concepts and Following Directions and Recalling Sentences, subtests of CELF-3. ^lComposite z score of narrative comprehension and production tasks, subtests of QRI-3.

* $p < .01$. ** $p < .001$.

right varied by rotation of 0°, 60°, or 120° clockwise from its original orientation. If the target figure was a mirror image, children were asked to press a different key.

5. The picture-matching task (PM) involved children deciding whether two pictures presented were the same or different. They were to press one key for “same” and another for “different.” In the first condition, the child decided whether the two pictures had the same physical shape; in the second, whether the pictures shared an identical name (e.g., two different cats); and in the third, whether the pictures were in the same category (e.g., animals).

Linguistic tasks. Five linguistic RT tasks were completed. In all of the linguistic tasks except picture naming, a positive response (stimulus is good, correct, or a match) was indicated by pressing a key with a green dot, and a negative response was indicated by pressing a key with a red dot.

1. In the picture-naming task (PN), children were asked to name pictures as soon as the stimulus was displayed on the screen. The names varied as to frequency of occurrence in English.
2. The truth value judgment task (TV) required children to determine whether an auditorily presented sentence did or did not correctly represent what a picture displayed.
3. The grammaticality judgment task (GJ) involved children listening to a sentence and deciding whether it was “good” or “bad.”
4. In the rhyme judgment task (RJ), the children decided whether two words rhymed or not. The children saw a picture depicting the first word at the top of the computer screen and judged whether it rhymed with a word presented 4 s later in an auditory, printed, or pictorial form.
5. The initial-consonant judgment task (IC) followed the same format as the RJ task and required children to decide whether the initial consonant of a picture name matched the initial consonant of the word presented next.

Data trimming. The data were trimmed using a procedure similar to previous RT studies (see C. Miller et al., 2001, 2006). Only correct responses were included, and accuracy was high across the tasks, the groups, and the grades (range of accuracy: 82% to 98%). The RTs were converted to z scores within each condition, so that all tasks were on a comparable scale.

Analyses

Diagnostic accuracy. A ROC curve was generated to define the best cutoff on each measure for each grade, third and eighth. The best cutoff for each measure was determined at a point on the ROC curve, which maximized classification

accuracy—that is, where sensitivity plus specificity divided by two is largest (Sackett et al., 1991). Using these cutoffs, likelihood ratios were computed for each of the tasks as measures of diagnostic accuracy. The LR+ was calculated as the ratio of true positives to false positives (sensitivity/[1-specificity]). The LR– was calculated as the ratio of false negatives to true negatives ([1-sensitivity]/specificity). Likelihood ratios and cutoffs are shown in Table 2.

Logistic regression analysis. A binary logistic regression analysis was used to examine whether any combination of processing speed measures would increase accuracy in predicting children’s language status more than any single measure. Diagnostic accuracy is addressed by how well a logistic regression model predicts group membership with given predictors in the model.

The processing speed measures were entered as independent variables sequentially, one by one. The entering order was determined by LR+ and LR– of the individual tasks, starting with the task with the best LR values. To assess if adding a new predictor to the model increased diagnostic accuracy, the likelihood ratio tests and associated χ^2 values were examined to determine significance of each predictor. The odds ratios were estimates of how much the probability of the categories of group membership (LI or TD) changed when the value of a predictor increased by one unit. Thus, odds ratios can be interpreted like effect sizes (Tabachnick & Fidell, 2013).

Results

Diagnostic Accuracy

The best cutoffs determined by ROC curves and LR+ and LR– values for the individual and combined processing speed measures are presented in Table 2. None of the individual or combined processing speed measures reached conventional levels of diagnostic accuracy as indicated by LR+ value of 10 or more and LR– value of 0.1 or less. In Grade 3, the GJ task had LR+ and LR– values of 6.29 and 0.61, respectively. A measure combining all processing speed tasks across domains provided poorer LR+ (3.05) but slightly better LR– (0.55) values than the GJ task alone. In Grade 8, the RJ task had LR+ and LR– values of 4.59 and 0.44, respectively. Similar to Grade 3, the overall combined measure of processing speed provided poorer LR+ (3.96) but slightly better LR– (0.40) values than the RJ task alone.

Logistic Regression

We then used logistic regression analyses to determine whether any combination of the processing speed tasks could offer better classification accuracy.

Grade 3. Table 3 summarizes the results of the logistic regression for Grade 3. Adding the GJ task (the predictor with the best diagnostic accuracy) to the constant only model significantly improved the model fit ($\chi^2 [df = 1, n = 130] = 16.67, p < .001$). Adding the RJ task (the predictor with the second best diagnostic accuracy) did not

Table 2. For Grades 3 and 8, z score cutoffs, positive likelihood ratios, and negative likelihood ratios on individual processing measures in order of diagnostic accuracy.

Grade 3								Grade 8							
Task	Cutoff	LR+	CI		LR–	CI		Task	Cutoff	LR+	CI		LR–	CI	
			L	U		L	U				L	U		L	U
GJ	0.96	6.29	2.69	14.67	0.61	0.47	0.80	RJ	0.31	4.59	2.58	8.18	0.44	0.30	0.65
RJ	–0.12	2.19	1.59	3.01	0.33	0.19	0.60	TV	0.71	4.47	2.21	9.04	0.61	0.46	0.80
IC	–0.08	2.12	1.46	3.08	0.49	0.31	0.75	SR	–0.38	3.39	1.99	5.76	0.52	0.36	0.73
PN	–0.31	1.70	1.31	2.20	0.33	0.17	0.67	PM	0.47	2.87	1.72	4.77	0.56	0.40	0.78
MR	0.07	1.74	1.15	2.63	0.66	0.47	0.93	IC	–0.14	2.34	1.66	3.30	0.35	0.20	0.61
TV	–0.14	1.74	1.25	2.42	0.52	0.33	0.83	VS	0.02	2.38	1.58	3.59	0.49	0.32	0.74
VS	–0.23	1.49	1.13	1.97	0.52	0.31	0.88	MR	–0.18	2.05	1.49	2.83	0.38	0.22	0.66
TP	–0.25	1.48	1.11	1.97	0.55	0.34	0.91	TP	0.12	2.12	1.38	3.26	0.59	0.41	0.84
SR	–0.09	1.44	1.02	2.03	0.66	0.44	1.01	GJ	–0.22	1.93	1.40	2.66	0.42	0.25	0.71
PM	–0.47	1.26	0.98	1.62	0.63	0.37	1.08	PN	–0.52	1.46	1.16	1.85	0.40	0.20	0.81
AP	0.36	3.05	1.81	5.13	0.55	0.39	0.77	AP	0.41	3.96	2.38	6.58	0.40	0.26	0.62
NP	0.55	2.37	1.38	4.07	0.68	0.51	0.90	NP	–0.04	2.52	1.74	3.65	0.36	0.22	0.62
LP	–0.57	1.62	1.33	1.98	0.17	0.06	0.51	LP	0.04	2.42	1.71	3.44	0.35	0.20	0.60

Note. LR+ = positive likelihood ratio; LR– = negative likelihood ratio; CI = 95% confidence interval; L = lower; U = upper; GJ = grammaticality judgment; RJ = rhyme judgment; IC = initial consonants; PN = picture naming; MR = mental rotation; TV = truth value judgment; VS = visual search; TP = tapping; SR = simple response; PM = picture matching; AP = combined measure of all processing tasks; NP = combined measure of nonlinguistic processing speed measures; LP = combined measure of linguistic processing speed measures.

significantly improve the model fit ($\chi^2 [1, n = 130] = 1.70$, $p = .192$). Thus, the RJ task was removed from the model. We then examined all the rest of the predictors one by one, but none significantly improved the model fit, and they were consequently removed from the model. Therefore, the final model with the best fit contained the GJ task as a single predictor of group membership. This model explained 16.8% of variance (Nagelkerke's R^2) in group membership. At the best cutoff (0.49) based on the ROC curve, it yielded an LR+ value of 6.29 and an LR– value of 0.61. The

LR+ value above 3.0 and below 10 indicated that slow grammaticality judgments were moderately suggestive of identifying LI, whereas the LR– value above .30 indicated that fast grammatical judgments were inadequate for ruling out LI.

Grade 8. Table 3 shows the results of the logistic regression for Grade 8. Adding the RJ task (the predictor with the best diagnostic accuracy) to the constant only model significantly improved the model fit ($\chi^2 [df = 1, n = 131] = 30.00$, $p < .001$). Second, the TV task, the

Table 3. Binary logistic regression models predicting probability of having language impairment (LI) in Grades 3 & 8.

Grades	Predictors		B	SE	Odds ratios	95% CI for odds ratios		χ^2
						Lower	Upper	
3	Step 0	Const	–0.74	0.19	0.48			
	Step 1 (final)	GJ	0.80	0.21	2.22	1.47	3.36	16.67**
		Const	–0.82	0.20	0.44			
	Step 2	GJ	0.59	0.25	1.81	1.11	2.96	
		RJ	0.39	0.30	1.48	0.82	2.66	1.70
8	Step 0	Const	–0.82	0.21	0.44			
		Const	–0.75	0.19	0.47			
	Step 1	RJ	1.18	0.26	3.27	1.97	5.42	30.00**
		Const	–0.87	0.22	0.42			
	Step 2	RJ	0.96	0.29	2.61	1.48	4.60	2.44
		TV	0.40	0.26	1.49	0.90	2.48	
	Step 3 (final)	Const	–0.90	0.22	0.41			
		RJ	1.04	0.28	2.83	1.65	4.85	
		SR	1.02	0.31	2.76	1.50	5.10	11.69*
		Const	–0.90	0.23	0.41			

Note. Positive coefficients for processing speed tasks indicate that slow processing speed is likely to predict children as having LI. SE = standard error; CI = confidence interval; Const = constant; final = final model selected (classification cutoff: 0.5); GJ = grammaticality judgment; RJ = rhyme judgment; TV = true value judgment; SR = simple response time tasks. All predictors are standardized.

* $p < .01$. ** $p < .001$.

predictor with the second best diagnostic accuracy among the predictors, was entered in the model. Adding the TV task did not improve the model fit ($\chi^2 [1, n = 131] = 2.44, p = .118$). Thus, it was removed from the model. Third, the SR task was entered in the previous model containing the constant and the RJ task. Adding the SR task significantly improved the model fit ($\chi^2 [1, n = 131] = 11.69, p < .001$). Thus, SR was kept in the model. We then examined all the rest of the predictors one by one, but none of them significantly improved the model fit, and they were consequently removed from the model. Therefore, the final model with the best fit included the RJ and SR tasks as a combination of predictors of group membership. At the best cutoff (0.52) based on the ROC curve, this composite model explained 37.7% of the variance in group membership (Nagelkerke's R^2). It yielded an LR+ value of 13.02 and an LR- value of 0.43. The LR+ value above 10 indicated that slow rhyme judgment together with slow simple response time were accurate in identifying LI. However, the LR- value was higher than the moderately suggestive value of 0.30, which indicated that fast responses on RJ and SR tasks were inadequate for ruling out LI.

Discussion

The current study investigated the relative utility of linguistic and nonlinguistic processing speed tasks as predictors of LI in children across two time points. We found that some processing speed measures were moderately informative but more effective in predicting the presence of LI than its absence. The results showed that in Grade 3, a single linguistic processing measure, the grammaticality judgment task (LR+ of 6.29 and LR- of 0.61) yielded the best diagnostic accuracy. However, in Grade 8, a combination of nonlinguistic (the simple response time task) and linguistic processing measures (the RJ task) resulted in the best diagnostic accuracy (LR+ of 13.02 and LR- of 0.43). Although this combined measure did not reach an excellent level of diagnostic accuracy, adding the nonlinguistic component to the linguistic component significantly enhanced diagnostic accuracy, and the proportion of variance in group membership accounted for increased from 28.6% to 37.7%. This indicates that both linguistic and nonlinguistic processing measures are associated with the diagnosis of LI, but only at Grade 8, when the participants were early adolescents.

The finding that different measures were useful for Grades 3 and 8 highlights the importance of considering a developmental perspective when developing better diagnostic tools. The measures with the best diagnostic accuracy among a set of 10 changed from the grammaticality judgment task in Grade 3 to a combination of rhyming judgment and simple response time tasks in Grade 8. The emergence of the simple response time task as a significant predictor in Grade 8 may indicate that deficits in nonverbal processes (perhaps motor with a timing component) become more apparent over time. When TD children are nearing the peak of processing speed (Kail, 1986), children with LI may have persistently delayed development of processing speed; thus,

in Grade 8, the gap between children with LI and TD peers became greater than in Grade 3, as the LR+ and LR- values suggest. Also, the proportion of variance in group membership accounted for increased from 16.8% to 37.7%, suggesting that processing speed is more consistently associated with LI in early adolescence compared to middle childhood.

Limitations of this study include the fact that it was a retrospective study of archival data. The processing speed tasks were not selected to address questions of diagnostic accuracy. The tasks were, however, selected to sample a range of stimuli and skills. We believe it is appropriate to use this unique data set to address our research questions. It is possible that the findings are unique to the specific tasks that were used. We make no claims about the relation of specific tasks to diagnostic status. Rather, we consider this a preliminary demonstration that processing speed tasks in general may have moderate value in predicting LI status and that nonlinguistic tasks may become more predictive later in development. However, we were only able to sample two time points in development. Future research should investigate whether other processing speed measures can be useful at an earlier age and what processing speed measures would be most predictive of LI at different ages. In addition, the data analyses in the current study included only children whose group classification (LI vs. TD) did not change over time, in order to maximize accurate classification. However, this procedure may tend to restrict the groups to participants with more extreme language scores, which could have an impact on the current results.

The literature on processing speed in children with LI indicates that these children are slower than same-age peers on a variety of tasks. There has been mixed evidence regarding the consistency of slowing across different types of tasks (e.g., compare C. Miller et al., 2006, and Windsor, Milbrath, Carney, & Rakowski, 2001) and regarding the predictive value of processing speed measures for language outcomes. We used a diagnostic accuracy approach to appraise the value of linguistic and nonlinguistic processing speed measures in predicting LI versus TD group membership. We found that at the earlier time point, in middle childhood, only a linguistic measure of grammaticality judgment emerged as a significant predictor of group status, and it was more effective at ruling in LI than ruling it out. At the later time point, in early adolescence, the combination of a simple nonlinguistic processing speed task and a linguistic task tapping phonological awareness provided the best prediction of group status and, again, were more effective at ruling in LI than ruling it out. This difference, with better LR+ values than LR- values, suggests that fast processing speed is observed in almost all TD individuals and in some individuals with LI but that slow processing speed is rarely observed in TD individuals. Thus, slow processing speed may be a good indicator of the presence of LI, especially in adolescents. Clinically, this study emphasizes the need to take developmental change into account when selecting diagnostic tools. In addition, these results

suggest that nonlinguistic processing could have a role to play in the identification of LI in older children.

Acknowledgments

The original study was supported by a clinical research center Grant P50 DC002746 from the National Institute on Deafness and Other Communication Disorders, awarded to J. Bruce Tomblin. We thank the Child Language Research Center at the University of Iowa and members of the Midwest Collaboration on SLI for the use of data. A preliminary report of these data was presented at the Symposium on Research in Child Language Disorders, Madison, Wisconsin, in June 2013 and at the American Speech-Language-Hearing Association Conference in Chicago, IL, in November 2013.

References

- Clegg, J., Hollis, C., Mawhood, L., & Rutter, M. (2005). Developmental language disorders—a follow-up in later adult life. Cognitive, language and psychological outcomes. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 46, 128–149.
- Dollaghan, C. A. (2007). *The handbook for evidence-based practice in communication disorders*. Baltimore, MD: Brookes.
- Ehren, B. J. (2002). Vocabulary intervention to improve reading comprehension for students with learning disabilities. *Perspectives on Language Learning and Education*, 9, 12–18.
- Fey, M. E., Catts, H. W., & Proctor-Williams, K. (2000, November). *Narrative generation by school-age children with typical and impaired language*. Presentation at the Annual Conference of the American Speech-Language-Hearing Association, Washington, DC.
- Johnson, C. J., Beitchman, J. H., Young, A. R., Escobar, M., Atkinson, L., Wilson, B., & Lam, I. (1999). Fourteen year follow-up of children with and without speech/language impairments: Speech/language stability and outcomes. *Journal of Speech, Language, and Hearing Research*, 42, 744–760.
- Kail, R. (1986). Sources of age differences in speed of processing. *Child Development*, 57, 969–987.
- Kail, R. V., & Miller, C. A. (2006). Developmental change in processing speed: Domain specificity and stability during childhood and adolescence. *Journal of Cognition and Development*, 7, 119–137.
- Kohnert, K., & Windsor, J. (2004). The search for common ground: Part II. Nonlinguistic performance by linguistically diverse learners. *Journal of Speech, Language, and Hearing Research*, 47, 891–903.
- Lahey, M., Edwards, J., & Munson, B. (2001). Is processing speed related to severity of language impairment? *Journal of Speech, Language, and Hearing Research*, 44, 1354–1361.
- Leonard, L. B. (2014). *Children with specific language impairment* (2nd ed.). Cambridge, MA: MIT Press.
- 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, 408–428.
- Miller, C., Kail, R., Leonard, L., & Tomblin, J. B. (2001). Speed of processing in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, 44, 416–433.
- Miller, C., Leonard, L., Kail, R., Zhang, X., Tomblin, B., & Francis, D. (2006). Response time in 14-year-olds with language impairment. *Journal of Speech, Language, and Hearing Research*, 49, 712–728.
- Miller, L. T., & Vernon, P. A. (1996). Intelligence, reaction time and working memory in 4- to 6-year-old children. *Intelligence*, 22, 155–190.
- Reilly, S., Tomblin, B., Law, J., McKean, C., Mensah, F. K., Morgan, A., . . . Wake, M. (2014). Specific language impairment: A convenient label for whom? *International Journal of Language & Communication Disorders*, 49, 416–451.
- Sackett, D. L., Haynes, R. B., Guyatt, G. H., & Tugwell, P. (1991). *Clinical epidemiology: A basic science for clinical medicine*. Boston, MA: Little, Brown.
- Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed.). Boston, MA: Pearson Education.
- Tomblin, J. B. (2010). The EpiSLI Database: A publicly available database on speech and language. *Language, Speech, and Hearing Services in Schools*, 41, 108–117.
- Tomblin, J. B., Records, N. L., Buckwalter, P., Zhang, X., Smith, E., & O'Brien, M. (1997). Prevalence of specific language impairment in kindergarten children. *Journal of Speech, Language, and Hearing Research*, 40, 1245–1261.
- Tomblin, J. B., Records, N., & Zhang, X. (1996). A system for the diagnosis of specific language impairment in kindergarten children. *Journal of Speech, Language, and Hearing Research*, 39, 1284–1294.
- Tomblin, J. B., Zhang, X., Buckwalter, P., & Catts, H. (2000). The association of reading disability, behavioral disorders, and specific language impairment in second grade children. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 41, 473–482.
- Wechsler, D. (1991). *Wechsler Intelligence Scale for Children—III*. San Antonio, TX: The Psychological Corporation.
- Whitehouse, A. J. O., Watt, H. J., Line, E. A., & Bishop, D. V. M. (2009). Adult psychosocial outcomes of children with specific language impairment, pragmatic language impairment and autism. *International Journal of Language & Communication Disorders*, 44, 511–528.
- Windsor, J., Kohnert, K., Loxtercamp, A., & Kan, P. F. (2008). Performance on nonlinguistic visual tasks by children with language impairment. *Applied Psycholinguistics*, 29, 237–268.
- 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, 446–461.
- Zhang, X., & Tomblin, J. B. (2000). The association of intervention receipt with speech-language profiles and social demographic variables. *American Journal of Speech-Language Pathology*, 9, 345–357.