ASHA American Speech-Language-Hearing Association

JSLHR

Research Article

Grammaticality Judgments of Tense and Agreement by Children With and Without Developmental Language Disorder Across Dialects of English

Janna B. Oetting,[a] Janet L. McDonald,[a] and Lori E. Vaughn[a]

[a] Louisiana State University, Baton Rouge

ARTICLE INFO

Article History:
Received March 19, 2023
Revision received June 29, 2023
Accepted September 6, 2023

Editor-in-Chief: Julie A. Washington
Editor: Mahchid Namazi

https://doi.org/10.1044/2023_JSLHR-23-00183

ABSTRACT

Purpose: Within General American English (GAE), the grammar weaknesses of children with developmental language disorder (DLD) have been documented with many tasks, including grammaticality judgments. Recently, Vaughn et al. replicated this finding with a judgment task targeting tense and agreement (T/A) structures for children who spoke African American English (AAE), a dialect that contains a greater variety of T/A forms than GAE. In the current study, we further tested this finding for children who spoke Southern White English (SWE), another dialect that contains a greater variety of T/A forms than GAE but less variety than AAE. Then, combining the SWE and AAE data, we explored the effects of a child's dialect, clinical group, and production of T/A forms on the children's judgments.
Method: The data were from 88 SWE-speaking children (DLD, *n* = 18; typically developing [TD], *n* = 70) and 91 AAE-speaking children (DLD, *n* = 34; TD, *n* = 57) previously studied. As in the AAE study, the SWE judgment data were examined both with *A′* scores and percentages of acceptability, with comparisons between dialects made on percentages of acceptability.
Results: As in AAE, the SWE DLD group had significantly different *A′* scores and percentages of acceptability than the SWE TD group for all sentence types, including those with T/A structures. Additional analyses indicated that the judgments of the TD but not the DLD groups showed dialect effects. Except for verbal *–s*, overt production and grammaticality judgments were correlated for the TD but not for the DLD groups.
Conclusions: Children with DLD across dialects of English present grammar difficulties that affect their ability to make judgments about sentences. More cross-dialectal research is needed to better understand the grammatical weaknesses of childhood DLD, especially for structures such as verbal *–s* that are expressed differently across dialects of English.

Like many other measures, grammaticality judgment tasks have helped reveal the grammar weaknesses of children with developmental language disorder (DLD; Leonard et al., 2009; Miller et al., 2009; Noonan et al., 2014; Rice et al., 1999, 2009).[1] To demonstrate competency on these tasks and as illustrated in 1a and 1b, an individual must judge grammatical sentences as *good* and ungrammatical sentences as *not so good*.

1a. Grammatical: *He is a dog*. Judgment: *Good*

1b. Ungrammatical: *He is run*. Judgment: *Not so good*

Correspondence to Janna B. Oetting: cdjanna@lsu.edu. ***Disclosure:*** *The authors have declared that no competing financial or nonfinancial interests existed at the time of publication.*

[1] We adopt the term *DLD* rather than specific language impairment (SLI) to promote uniformity across researchers and public awareness of the childhood language condition while also noting that our inclusionary and exclusionary criteria for DLD are consistent with previous studies that select children with SLI from a larger and more heterogeneous group of children with DLD.

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

A key assumption of grammaticality judgment tasks is that the sentences classified by the examiner as grammatical and ungrammatical are also grammatical and ungrammatical in the dialects of the individuals tested (Garcia et al., 2022). For children learning General American English (GAE),[2] this assumption is often met because the tasks are typically created using GAE as the dialect reference. Unfortunately, this assumption is often not met when children are learning a dialect of English that differs from GAE.

There is good reason to be cautious about administering judgment tasks and other types of tasks created for GAE to evaluate the language abilities of children learning other dialects of English, such as African American English (AAE). Following the seminal work of Johnson (2005) and de Villiers and Johnson (2007), Beyer and Hudson Kam (2012) administered a GAE-based picture choice task to typically developing (TD) first and second graders learning either AAE or GAE. The task presented one sentence and two sets of pictures to the children and asked them to select the picture set that best matched the sentence. The authors described their sentences as one of two types: unshared and shared. The unshared sentences contained grammatical content that the authors described as occurring infrequently in AAE but frequently in GAE (i.e., overt forms of past tense, such as *walked*; future modals, such as *'ll*; and verbal *–s*, such as *walks*). The shared sentences were the same as the unshared, except these sentences also had grammatical content shared across AAE and GAE (i.e., temporal words, such as *yesterday*, and overt plural forms, such as *cats*) that could be used to select the pictures.

Beyer and Hudson Kam's (2012) results varied by sentence type. For the shared sentences, no dialectal differences were detected, and both dialect groups performed above chance. For the unshared sentences, dialectal differences surfaced. Specifically, the AAE group did not perform above chance, whereas the GAE group performed above chance on the unshared sentences with overt past tense forms and modals (and the GAE second graders were also above chance on sentences with overt verbal *–s* forms as gleaned from Figure 2, p. 375). These findings underscore the need to learn more about the judgments of children who speak AAE and other dialects of English that differ from GAE, as well as learn how children with DLD across various dialects perform on judgment tasks relative to their same dialect-speaking TD peers.

[2]We refer to all dialects of English in the United States as GAE when they are classified by an author as standard and/or mainstream, or when they are classified as English without noting the dialect of the participants. This practice seeks to name GAE, AAE, and SWE using conceptually equivalent labels, with the terms *mainstream English* and *nonmainstream English* referring to broader categories of dialects spoken in Australia, Britain, the Americas, and elsewhere.

The Grammatical Profile of Childhood DLD Across Dialects of English

In a series of studies, Oetting and colleagues have examined the grammatical profile of childhood DLD within AAE and Southern White English (SWE), two dialects spoken in the rural south that differ from GAE. These studies have repeatedly found children with DLD to present grammar weaknesses within their dialects (e.g., Oetting et al., 2019, 2021; Oetting & McDonald, 2001; Oetting, McDonald, et al., 2016; Rivière et al., 2018; for earlier work, see Oetting, 2019). For example, AAE- and SWE-speaking children with DLD have been shown to produce lower percentages of overt forms for several tense and agreement (T/A) structures (i.e., auxiliary BE as in *is, are, was, were*; past tense as in *mowed* and *ate*; verbal *–s* as in *glues*) as compared to same dialect-speaking TD controls. This same finding has been documented within GAE (Leonard, 2017; Rice, 2003; Rice et al., 1998). Across AAE, SWE, and GAE, children with DLD also have not been found to produce high rates of overt T/A forms in language samples in contexts where these forms are ungrammatical within their dialects (Oetting et al., 2021, 2022; Rice & Wexler, 1996; Rice et al., 1995).

Grammaticality Judgments by AAE-Speaking Children With and Without DLD

Recently, Vaughn et al. (2023) extended the study of childhood DLD within AAE using a grammaticality judgment task. The children's judgments were collected using items from the grammaticality judgment subtest of the Test of Early Grammatical Impairment (TEGI; Rice & Wexler, 2001). As a test of GAE, the TEGI norms were not considered valid for AAE. Indeed, 47%–63% of the AAE-speaking TD children in this study would have failed the TEGI had the GAE normative data been considered. Nevertheless, we reasoned that the TEGI sentences could be used to learn about childhood DLD within AAE if the children's judgments were examined using a cross-linguistic (i.e., cross-dialectal) framework. To do this, the TEGI sentences were classified as one of four types based on their grammaticality in GAE and AAE: GAE- and AAE-appropriate T/A overt forms (e.g., *He is a bear, He looks big*), GAE-inappropriate but AAE-appropriate T/A zero forms (e.g., *He Ø hiding, Now the bear wantØ drink*), GAE- and AAE-inappropriate zero progressive forms (e.g., *He is buildØ*), and GAE- and AAE-inappropriate misapplications of overt T/A forms (e.g., *He am. . .*).

We also explored two scoring systems for the children's judgments. The first involved calculating *A′* scores based on GAE and the TEGI manual. *A′* scores consider the children's hits (i.e., judging grammatical sentences as

"good") and false alarms (i.e., judging ungrammatical sentences as "good"). For all A' scores, the GAE- and AAE-appropriate T/A overt forms were classified as grammatical, and all others were classified as GAE ungrammatical, even though the zero T/A forms were grammatical in AAE. Given this, we predicted that no clinical group differences would be found for the A' scores involving the zero T/A forms. Then, the children's responses were examined as percentages of acceptability based on AAE. With this system, percentages of the children's *good* responses were calculated for each form type, with the overt T/A forms and zero T/A forms treated as equivalent in grammaticality (Oetting et al., 2022; Shin, 2022; Shin & Miller, 2022).

Counter to our predictions, significant clinical group differences were found with both scoring systems; however, more dialect-specific information was learned about the children's language abilities using the acceptability percentages. For example, the DLD and TD groups did not differ on their acceptability of the zero T/A forms (as predicted), but the DLD group was less accepting of the T/A overt forms compared to the TD group, despite both groups judging these forms as acceptable at levels above chance. These results were not obtained with the A' scoring system because the overt and zero T/A forms were considered together to calculate the A' scores.

The analyses then turned to the children's judgments of *is* and verbal *–s*, two structures repeatedly targeted within the TEGI. Interestingly, both groups accepted the overt forms of *is* at higher levels than the zero forms of *is*, and their judgments of the zero forms negatively correlated with their production of these forms. In contrast, both groups accepted the overt forms and the zero forms for verbal *–s* at above chance levels, and their judgments of the zero forms were not correlated to their productions of these forms.

From these results, we concluded that judgment tasks that make use of acceptability percentages for specific types of overt forms and zero forms can be used to learn about childhood DLD within AAE and to discover important contrasts between T/A structures (e.g., *is* and verbal *–s*) within AAE. Extending the study of Vaughn et al. (2023) to other non-GAE dialects is needed to determine if the findings for T/A are specific to AAE or present in all dialects that make use of overt forms and zero forms.

Grammaticality Judgments by SWE-Speaking Children With and Without DLD

SWE provides an ideal test of Vaughn et al.'s (2023) findings because SWE speakers produce many of the same zero forms as AAE speakers, but they produce them at significantly lower frequencies (Cleveland & Oetting, 2013; Oetting et al., 2019, 2021; Oetting & Garrity, 2006; Oetting & McDonald, 2001). In addition, the AAE speakers studied by Vaughn et al. were drawn from a multiyear investigation that included SWE speakers, and a subset of the SWE and AAE groups studied here contributed data to three previous T/A production studies (Oetting et al., 2019, 2021; Oetting, McDonald, et al., 2016). Thus, we have production data that can be considered alongside the children's grammaticality judgments. Most relevant to the current study are findings from Oetting et al. (2019) because probes were used to elicit the children's productions of overt forms and zero forms for BE present (*is, are, Ø*) and verbal *–s* (*sees*/*see Ø*), two structures repeatedly targeted on the TEGI. In the current study, the children's productions of these forms on the elicitation probes are compared to their judgments of these same forms.

Recall that the AAE-speaking children studied by Vaughn et al. (2023) judged forms for *is* differently than forms for verbal *–s*. Interestingly, the AAE- and SWE-speaking children studied by Oetting et al. (2019) also produced different patterns of results for BE present and verbal *–s* on the elicitation probes. Specifically, no dialectal differences were detected for the overt BE present forms for either the TD (AAE, 78% vs. SWE, 86%), $F(1, 51) = 1.06$, $p = .31$, $\eta_p^2 = .02$, or DLD (AAE, 53% vs. SWE, 53%), $F(1, 51) = 0.001$, $p = .98$, $\eta_p^2 < .001$, group; however, for the overt verbal *–s* forms, a dialectal difference was detected for the TD group but not for the DLD group (TD AAE, 46% vs. TD SWE, 92%, $F(1, 51) = 24.84$, $p < .001$, $\eta_p^2 = .33$; DLD AAE, 21% vs. DLD SWE, 37%, $F(1, 51) = 2.90$, $p = .095$, $\eta_p^2 < .05$). Comparing within each dialect, percentages of overt BE and overt verbal *–s* forms also differed for the AAE TD group (78% vs. 46%, respectively), $F(1, 34) = 30.86$, $p < .001$, $\eta_p^2 = .48$, but not for the SWE TD group (86% vs. 92%, respectively), $F(1, 17) = 1.47$, $p = .24$, $\eta_p^2 = .08$.[3] These findings reveal important dialectal differences in the children's productions of verbal *–s* overt forms (and inversely their zero forms), at least for children classified as TD. Whether these dialectal differences are also present when verbal *–s* forms are tested using a grammaticality judgment task is unknown.

Finally, it is important to ask whether the profile of childhood DLD across dialects varies when children are given a grammaticality judgment task as compared to a production task. The source of each type of difference

[3]Oetting et al. (2019) used three different scoring systems to calculate percentages of overt form use. Here, we present the 2019 findings using the dialect-informed, strategic scoring system and use the same scoring system in the current study.

(dialectal vs. DLD profile) is not the same. Dialectal differences reflect typical, sociolinguistic processes inherent to all language learning communities, whereas differences related to childhood DLD reflect atypical, psycholinguistic processes that affect a relatively small subset of learners within every language community (Oetting, 2019; Oetting, Gregory, & Rivière, 2016). Given this, we may expect differential effects of task manipulations, with perhaps greater task effects (or different task effects) observed for TD children whose dialects differ compared to those observed for children with DLD whose dialects differ. If observed, this pattern of findings would help motivate the development of theories that seek to explain typical and atypical language variation within the same model. To these authors' knowledge, a comprehensive theoretical model that includes task effects is lacking across fields.

Research Questions

1. Do children with and without DLD differ in their grammaticality judgments within SWE as has been found for AAE?
2. Considering SWE and AAE child data together, are there dialectal differences in children's grammaticality judgments—particularly for verbal *–s* forms that are produced differently across the two dialects?
3. Are there dialectal differences (SWE vs. AAE) and/or clinical group differences (DLD vs. TD) in relations between the children's grammaticality judgments and their T/A form productions?

Method

Participants

The data came from 179 kindergartners who lived in a rural southern community and were classified by dialect as a speaker of either SWE (n = 88) or AAE (n = 91) and by clinical group as either DLD (n = 52 total, n = 18 SWE, n = 34 AAE) or TD (n = 127 total, n = 70 SWE, n = 57 AAE). As previously mentioned, the judgments were collected as part of a multiyear investigation of childhood DLD, the judgment data of 91 AAE speakers were previously studied by Vaughn et al. (2023), and the production data of 36 SWE and 70 AAE speakers were previously studied by Oetting et al. (2019). Prior to data collection, the study was approved by the institutional review board at Louisiana State University A&M College, and parental consent and child assent were obtained for all participants. All children who met the inclusionary and exclusionary criteria for the previous studies (n = 180; Rivière et al., 2018) and had judgment data within their files (n = 179) were included in the analyses.[4]

Details related to the measures used to classify the children by dialect and clinical group can be found in the studies of Oetting, McDonald, et al. (2016), and Oetting et al. (2019, 2022). For convenience, a summary of the groups is presented here (see Table 1). All children passed a school-administered pure-tone hearing screening at 30 dB for 1000, 2000, and 4000 Hz in each ear; earned a nonverbal IQ standard score at or above 82 (−1.2 *SD*) on the Primary Test of Nonverbal Intelligence (PTONI; M = 100, SD = 15; Ehrler & McGhee, 2008); and earned a standard score at or above 85 (−1 *SD*) on the Goldman-Fristoe Test of Articulation–Second Edition (GFTA-2; M = 100, SD = 15; Goldman & Fristoe, 2000) Sounds-in-Words subtest. The GFTA-2 was scored without modification for dialect variation.

The children's dialects were confirmed as SWE or AAE using blinded listener judgments (Oetting & McDonald, 2002); the Diagnostic Evaluation of Language Variation–Screening Test (DELV-Screening Test; Seymour et al., 2003); and, in some cases, analysis of a language sample. The dialect judgments came from three trained graduate students who were native to Louisiana; each independently rated the children's dialects after listening to a short excerpt of conversational speech. During the task, the raters were blind to the children's race, gender, and clinical status (Oetting, McDonald, et al., 2016; Rivière et al., 2018). As expected and using a 7-point Likert scale (1 = *low*, 7 = *high*), the listeners rated the densities of the SWE group's dialect-specific form use lower than the AAE group's (SWE M = 2.97, SD = 1.19, range: 1.33–6.33 vs. AAE M = 4.7, SD = 1.32, range: 1.67–7.00). The SWE group's average percentage of dialect-specific forms on the DELV-Screening Test was also lower than the AAE group's average (SWE M = 48, SD = 29, range: 0–100 vs. AAE M = 82, SD = 22, range: 13–100).

The Syntax subtest of the Diagnostic Evaluation of Language Variation–Norm Referenced (DELV-NR; Seymour et al., 2005) was administered to classify the children as DLD or TD. Children classified as DLD earned a score of $\leq$ 7, and the children classified as TD earned a score of $\geq$ 8. The DELV-NR Syntax subtest has a normative mean of 10 (SD = 3, −1 SD = 7). Although the DELV-NR manual does not report diagnostic accuracy

[4]Unlike Oetting et al. (Oetting et al., 2019, 2021; Oetting, McDonald, et al., 2016) but consistent with Rivière et al. (2018), we allowed the dialect and clinical groups to be unmatched for nonverbal IQ following recent recommendations to allow a greater range of nonverbal IQ scores in studies of DLD (Bishop et al., 2017). This resulted in more children and greater heterogeneity in the TD groups than in our previous studies.

Table 1. Demographics and test results by dialect and clinical group.

Variable	SWE		AAE		Findings
	DLD *n* = 18	TD *n* = 70	DLD *n* = 34	TD *n* = 57	
Age[a]	65.72 (3.89)	66.77 (4.08)	67.12 (3.65)	65.00 (3.44)	Gp × Dialect: $F(1, 175) = 6.00$, $p = .02$, $\eta^2 = .03$
MED[b]	12.33 (2.87)	12.99 (2.63)	11.75 (2.26)	13.05 (2.58)	Gp: $F(1, 169) = 4.83$, $p = .03$, $\eta^2 = .03$
PTONI[c]	96.50 (8.35)	107.90 (16.51)	93.35 (9.55)	100.98 (11.92)	Gp: $F(1, 175) = 17.53$, $p < .001$, $\eta^2 = .09$ Dialect: $F(1, 175) = 4.91$, $p = .03$, $\eta^2 = .03$
GFTA-2[d]	104.78 (4.18)	109.11 (4.68)	104.35 (5.75)	107.33 (4.52)	Gp: $F(1, 175) = 19.84$, $p < .001$, $\eta^2 = .10$
DELV-NR Syntax[e]	4.78 (1.66)	10.17 (1.66)	4.85 (1.02)	9.84 (1.46)	Gp: $F(1, 175) = 413.21$, $p < .001$, $\eta^2 = .70$
PPVT-4[f]	85.78 (7.01)	107.23 (7.81)	82.32 (9.56)	101.46 (9.02)	Gp: $F(1, 175) = 188.97$, $p < .001$, $\eta^2 = .52$ Dialect: $F(1, 175) = 9.77$, $p = .002$, $\eta^2 = .05$

Note. Means reported with standard deviations in parentheses. SWE = Southern White English; AAE = African American English; DLD = developmental language disorder; TD = typically developing; Gp = Group; MED = Maternal Education Level.

[a]Reported in months. [b]Reported in years of schooling (i.e., 12 = high school graduate). [c]Standardized scores for the Primary Test of Nonverbal Intelligence (PTONI; normative $M = 100$, $SD = 15$). [d]Standardized scores for the Goldman-Fristoe Test of Articulation–Second Edition (GFTA-2; normative $M = 100$, $SD = 15$). [e]Standardized scores for the Syntax portion of the Diagnostic Evaluation of Language Variation–Norm Referenced (DELV-NR; normative $M = 10$, $SD = 3$). [f]Standardized scores for the Peabody Picture Vocabulary Test–Fourth Edition (PPVT-4; normative $M = 100$, $SD = 15$).

indices for subtests, the manual reports Se and Sp values of .95 and .93, respectively, when ≤ -1 *SD* is used as the cut score on the composite language score (p. 140).

The Peabody Picture Vocabulary Test–Fourth Edition (PPVT-4; Dunn & Dunn, 2007) was also administered as a descriptive measure of the children's vocabulary abilities. The PPVT-4 has a normative mean of 100 ($SD = 15$). Finally, 87 families of children classified as SWE completed the family history question on the consent form, and of those, 25 (29%) reported a positive family history of speech or language impairment. As found in many other DLD studies (Leonard, 2017; Pruitt et al., 2010), a positive family history was reported for more of the SWE DLD group than for the TD group (44% vs. 25%, respectively); this same finding was reported by Vaughn et al. (2023) for the AAE DLD and TD groups (44% vs. 18%, respectively).

Although the first research question focused on the SWE groups, the other two questions focused on dialect differences (SWE vs. AAE). Given this, the data in Table 1 were examined with 2 (dialect) × 2 (clinical group) analyses of variance. There was a Dialect × Clinical Group interaction for child age, which was related to differences (DLD > TD) in the AAE but not SWE group, and a main effect for clinical group (DLD < TD) for level of maternal education ($\eta_p^2 = .045$). Main effects for clinical group were also observed for the various tests administered, with the largest effect size observed for the DELV-NR Syntax subtest ($\eta_p^2 = .70$). Dialect differences were also detected for the PTONI and PPVT-4, although the effect sizes of these differences were small ($\eta_p^2 = .03$ and .05, respectively).

Grammaticality Judgment Task

The TEGI judgment task included 35 sentences; 10 with T/A overt forms—three of which also included three overt progressive forms, 10 with T/A zero forms, five with zero progressive forms, and 10 with overt T/A misapplications.[5] Also, all overt forms of BE (*am, is, are*) were presented within the sentences as uncontracted morphemes (to download items, see https://kuscholarworks.ku.edu/bitstream/handle/1808/32569/TEGI_ScoreSheetComplete.pdf?sequence=2&isAllowed=y). The TEGI grammaticality judgment task was administered in live voice following the

[5]The overt misapplications involved three *he am*, three *he are*, three *I likes/drinks/wants*, and one *you needs*. Our previous studies have shown these forms to be rarely produced by SWE and AAE child speakers. For example, within 5,231 utterances from 45 AAE child speakers, Garrity and Oetting (2010) identified 74 overt *am* forms, but none involved *he am*. Within 12,595 utterances from 57 AAE and SWE child speakers, Cleveland and Oetting (2013) identified 1,159 overt verbal *–s* forms, and only 14 (1%) involved a first-, second-, or third-person plural subject. Finally, within 25,106 utterances from some of the SWE and AAE child speakers studied here, Oetting et al. (2021) identified 1,689 verbal *–s* contexts and 665 *are* contexts. Of these, only two (< 1%) of the overt verbal *–s* forms were produced with a first-person subject, and none of the overt *are* forms were produced with a singular subject.

examiner's manual (Rice & Wexler, 2001). The task began with the examiner presenting the children a script involving Zee and Bo, two robots who were visiting earth and learning English. As part of the script and after each sentence with a target T/A structure, the children were prompted to judge the grammaticality of Zee or Bo's English.

The children's judgments were first converted to three A' scores based on GAE. A' scores consider hits (grammatical sentences for GAE judged as grammatical) and false alarms (ungrammatical sentences for GAE judged as grammatical); an A' score of .50 indicates chance performance, and a score of 1.00 indicates judgments that match the items' grammaticality in GAE. For each A' score, the GAE-appropriate overt T/A forms served as the grammatical items and the others as the ungrammatical items.

The children's judgments were then examined using percentages of acceptability (i.e., percentage of responses indicating that a sentence was *good*). These percentages were calculated separately for the T/A overt forms and T/A zero forms, along with separate calculations for the zero progressive forms and overt T/A misapplications. For these, chance was 50%; percentages significantly above this level were interpreted as accepted, and percentages significantly below this level were interpreted as rejected.

T/A Elicitation Probes

The elicitation probes included 64 short videos paired with verbal prompts to elicit 16 forms for auxiliary BE present (*is*, *are*, *Ø*), auxiliary BE past (*was*, *were*, *Ø*), main verb past tense (e.g., *mowed*, *mowØ*), and main verb verbal *–s* (*glues*, *glueØ*). Following the study of Vaughn et al. (2023), the children's percentages of overt form use were calculated using Oetting et al.'s (2019) strategic scoring approach (i.e., sum of all dialect appropriate overt forms divided by the sum of these forms and the child's zero forms).

Reliability

Reliability of the TEGI scoring was evaluated by Vaughn et al. (2023) by asking a second examiner to independently score 20% of the AAE-speaking children's judgments. The rate of agreement was 98%, which was considered sufficient for confirming the reliability of the AAE and SWE judgment data, as these were collected and scored at the same time and with the same examiners. Reliability of the elicitation probe data was checked by having a second set of examiners independently score 20% of the children's T/A productions as part of the multiyear investigation, which included the SWE and AAE data studied here. The average rate of agreement between the two sets of examiners was 95%.

Results

The first set of analyses included data from the SWE group only. As was found for AAE, the SWE-speaking children's judgments were not normally distributed (across measures, the skew of the data for the DLD group ranged from −1.076 to 0.157, and the skew of the data for the TD group ranged from −2.842 to 0.86). Therefore, following the study of Vaughn et al. (2023), Mann–Whitney U tests were used to examine between-groups differences, Wilcoxon signed-rank tests were used to test the children's judgments against chance, and Friedman's tests were used to examine within-group differences. Then, the data from the SWE and AAE groups were combined and examined with additional Mann–Whitney U tests and Spearman correlations. This approach allowed us to test the previous AAE findings within SWE as well as examine dialect and clinical group differences in the children's judgments.

SWE-Speaking Children's Judgments

Table 2 reports the SWE-speaking children's judgment data and statistical results for the A' and the percentages of acceptability measures. As was found for AAE, judgments by the SWE DLD and TD groups differed from each other regardless of whether A' scores or percentages of acceptability were examined.

A' Scores

Recall that for the A' scores, the T/A overt forms (e.g., *He is a bear*) served as the GAE grammatical items, and there were three types of GAE ungrammatical structures: T/A zero forms (e.g., *And he Ø furry*), zero progressives forms (e.g., *He is buildØ*), and overt T/A misapplications (e.g., *He am…*). This resulted in three A' scores. Results indicated that all the DLD group's A' scores were lower than the TD group's A' scores. In addition, none of the SWE DLD A' scores were above chance (zero T/A forms, $Z = 0.40$, $p = .69$; zero progressive forms, $Z = 1.03$, $p = .30$; overt T/A misapplications, $Z = 1.07$, $p = .28$), whereas all three of the TD A' scores were (zero T/A forms, $Z = 7.26$, $p < .001$; zero progressive forms, $Z = 7.41$, $p < .001$; overt T/A misapplications, $Z = 7.27$, $p < .001$). Finally, Friedman's tests indicated that there was not a significant difference between the three A' scores of the DLD group, $\chi^2(2) = 2.21$, $p = .33$, $W = .005$, whereas there was for the TD group, $\chi^2(2) = 43.03$, $p < .001$, $W = .94$, with a Bonferroni-corrected Dunn post hoc test showing the TD group's scores for the T/A zero forms to be lower than their other scores.

Table 2. Southern White English *A′* values and percentages of acceptability by grammatical category and clinical group.

Grammatical category	DLD	TD	*U*	*Z*	*p*
GAE *A′* zero T/A forms	.51 (.22)	**.84**[a] (.13)	95.50	−5.54	< .001
GAE *A′* zero progressive forms	.56 (.27)	**.93**[b] (.14)	184.00	−4.97	< .001
GAE *A′* overt T/A misapplications	.56 (.25)	**.91**[b] (.11)	113.50	−.5.44	< .001
% Acceptable overt T/A forms	**68** (35)	**88**[a] (18)	394.50	−2.55	.011
% Acceptable zero T/A forms	59 (32)	**28**[b] (23)	279.50	3.65	< .001
% Acceptable zero progressive forms	48 (41)	**12**[c] (19)	304.00	3.70	< .001
% Acceptable overt T/A misapplications	54 (33)	**15**[c] (18)	191.50	4.65	< .001

Note. Means reported with standard deviations in parentheses. Bolded numbers are significantly different from chance. Different letters indicate within-group differences within the *A′* scores and within the percentages of acceptability. DLD = developmental language disorder; TD = typically developing; GAE = General American English; T/A = tense and agreement.

We also considered the GAE-based norms of the TEGI for the children's *A′* scores. Consistent with their clinical status, most of the children in the SWE DLD group failed the task (94%, 84%, and 94%, respectively). However, these norms also led to substantial numbers of children in the SWE TD group failing the task (zero T/A forms = 33%, zero progressive forms = 40%, and overt T/A misapplications = 33%), and the high fail rate of the TD group indicates that the TEGI norms are not appropriate for SWE.

Percentages of Acceptability

As also shown in Table 2, the SWE DLD group was less likely to judge the T/A overt forms as acceptable when compared to the SWE TD group, although both groups accepted these forms at levels above chance, DLD: $Z = 1.96$, $p = .05$, TD: $Z = 7.11$, $p < .001$. For the T/A zero forms, zero progressive forms, and overt T/A misapplications, the clinical groups also differed, but here the DLD group performed at chance (T/A zero forms, $Z = 1.17$, $p = .24$; zero progressive, $Z = 0.22$, $p = .82$; overt T/A misapplications, $Z = 0.55$, $p = .58$), and the TD group rejected them (T/A zero forms, $Z = -5.77$, $p < .001$; zero progressive, $Z = -7.10$, $p < .001$; overt T/A misapplications, $Z = -7.00$, $p < .001$). In addition, within-group comparisons indicated that the SWE DLD group's acceptability of the various forms did not differ from each other, but those of the SWE TD group did, $\chi^2(2) = 43.03$, $p < .001$, $W = .94$, with a Bonferroni-corrected Dunn post hoc test showing percentages for the overt T/A forms to be higher than percentages for the zero T/A forms and both of these to be higher than their percentages for the zero progressive forms and overt T/A misapplications.

To examine these data in more detail, the overt and zero forms of each T/A structure were examined. Of the 10 TEGI T/A overt forms, five were of *is*, three were of verbal *–s*, and two were of past tense. Of the 10 TEGI T/A zero forms, five were of *is* and five were of verbal *–s*. As shown in Table 3, the SWE DLD group was less likely to judge overt *is* forms as acceptable compared to the SWE TD group, but they did not differ in their acceptability of the overt verbal *–s* forms and overt past tense

Table 3. Southern White English percentages of acceptability by structure, form, and clinical group.

Structure and form	DLD	TD	*U*	*Z*	*p*
Overt *is* forms 5 items	62[b] (37)	**89**[a] (18)	324	−3.51	< .001
Overt verbal *–s* forms 3 items	**76**[a] (34)	**90**[a] (22)	506	−1.67	.096
Overt regular past tense forms 2 items	69 (43)	**82**[a] (34)	532	−1.30	.194
Zero *is* forms 5 items	49[b] (41)	**13**[b] (21)	313	3.65	< .001
Zero verbal *–s* forms 5 items	**69**[a] (29)	44[c] (30)	342	3.03	.002

Note. Means reported with standard deviations in parentheses. Bolded numbers are significantly different from chance. Different letters indicate within-group differences between forms. DLD = developmental language disorder; TD = typically developing.

forms (although these were in the direction of lower acceptability by the DLD group). The SWE DLD and TD groups also differed in their acceptability of the zero forms, with the DLD group accepting the zero *is* forms at chance, $Z = -0.18$, $p = .86$, and the zero verbal *–s* forms above chance, $Z = 2.36$, $p = .02$, and with the TD group rejecting the zero *is* forms, $Z = -7.05$, $p < .001$, while accepting the zero verbal *–s* forms at chance, $Z = -1.69$, $p = .09$. These results highlight structure-specific differences between the SWE DLD and TD groups' judgments.

In summary, SWE-speaking children with DLD were less able than their same dialect-speaking TD peers to judge the grammaticality of various types of sentences, including those with T/A overt forms and zero forms. Many of the observed clinical group differences within SWE were also found by Vaughn et al. (2023) for AAE, except for the children's acceptability percentages for the zero *is* forms and zero verbal *–s* forms. For these forms, the SWE DLD and TD groups differed, but the AAE DLD and TD groups did not. Whereas the *A′*s scores masked these dialectal differences, the acceptability percentages helped illuminate them. To learn more about these dialectal differences and perhaps others, a direct comparison of the SWE and AAE groups is needed.

Differences Between SWE and AAE

The acceptability percentages from the SWE and AAE groups are presented in Table 4. As shown, patterns of dialectal difference depended on the children's clinical status, with the TD groups, but not the DLD groups, presenting dialectal similarities and differences. Regarding similarities, the TD dialect groups did not differ in their acceptance of the overt *is* and overt verbal *–s* forms, and both groups accepted these forms at levels above chance (overt *is*: SWE: $Z = 7.31$, $p < .001$, AAE: $Z = 6.50$, $p < .001$; overt verbal *–s*: SWE: $Z = 7.16$, $p < .001$, AAE: $Z = 6.47$, $p < .001$). The TD groups also did not differ in their judgments of the zero progressive forms, and both groups rejected these forms at levels that differed from chance (SWE: $Z = -7.10$, $p < .001$; AAE: $Z = -4.66$, $p < .001$). Regarding differences, the SWE TD group's acceptability percentages for the zero *is* forms and overt T/A misapplications were lower than the AAE TD group's; however, like the zero progressive forms, both TD groups rejected these forms at levels that differed from chance (zero *is* forms: SWE: $Z = -7.05$, $p < .001$, AAE: $Z = -2.50$, $p = .013$; overt T/A misapplications: SWE: $Z = -7.00$, $p < .001$, AAE: $Z = -2.29$, $p = .02$). Finally, and most strikingly, a dialectal difference was observed for the zero verbal *–s* forms. Here, the SWE TD group performed at chance, $Z = -1.69$, $p = .09$, whereas the AAE TD group accepted these forms at levels above chance, $Z = 4.66$, $p < .001$.

For the children with DLD, dialect spoken had no influence on their percentages of acceptability. Not only were no dialectal differences detected, but both DLD groups accepted the overt verbal *–s* forms and zero verbal *–s* forms at levels above chance (overt verbal *–s*: SWE: $Z = 2.67$, $p = .01$, AAE: $Z = 2.50$, $p = .01$; zero verbal *–s*: SWE: $Z = 2.36$, $p = .02$, AAE: $Z = 3.68$, $p < .001$) and performed at chance for all others (overt *is* forms: SWE: $Z = 1.18$, $p = .24$, AAE: $Z = 1.48$, $p = .14$; zero *is* forms: SWE: $Z = -0.18$, $p = .86$, AAE: $Z = -0.14$, $p = .89$; zero progressive forms: SWE: $Z = -0.22$, $p = .82$, AAE: $Z = -1.63$, $p = .10$; overt TA misapplications: SWE: $Z = 0.55$, $p = .58$, AAE: $Z = 1.21$, $p = .23$).

Table 4. Mean acceptability percentages for various forms by clinical status and dialect.

	TD				DLD			
Form	**SWE**	**AAE**	***Z***	***p***	**SWE**	**AAE**	***Z***	***p***
Overt *is* forms	**89** (18)	**85** (19)	1.35	.18	62 (37)	59 (34)	0.42	.67
Overt verbal *–s* forms	**90** (22)	**87** (22)	1.08	.28	**76** (34)	**66** (33)	1.91	.23
Overt past tense forms	**82** (34)	**80** (34)	0.58	.56	69 (43)	60 (40)	0.88	.38
Zero *is* forms	**13** (21)	**38** (37)	–4.0	< .001	49 (41)	49 (31)	0.08	.94
Zero verbal *–s* forms	44 (30)	**72** (28)	–4.88	< .001	**69** (29)	**74** (29)	0.69	.49
Zero progressive forms	**12** (19)	**25** (32)	–1.89	.06	48 (41)	42 (28)	0.31	.75
Overt T/A misapplications	**15** (18)	**40** (31)	–5.01	< .001	54 (33)	56 (31)	–0.19	.85

Note. Means reported with standard deviations in parentheses. Bolded numbers are significantly different from chance. TD = typically developing; DLD = developmental language disorder; SWE = Southern White English; AAE = African American English; T/A = tense and agreement.

Table 5. Percentage of tense and agreement overt forms produced on the elicitation probes by clinical group and dialect.

Form	TD				DLD			
	SWE	AAE	*Z*	*p*	SWE	AAE	*Z*	*p*
Overall	91 (09)	69 (20)	6.24	< .001	48 (30)	42 (21)	0.65	.51
Auxiliary *is*	97 (08)	79 (32)	3.85	< .001	64 (38)	56 (36)	0.82	.41
Verbal *–s*	91 (14)	47 (36)	7.15	< .001	37 (40)	19 (22)	1.01	.31

Note. Means reported with standard deviations in parentheses. TD = typically developing; DLD = developmental language disorder; SWE = Southern White English; AAE = African American English.

We now turn to examining what might predict a speaker's acceptability judgments within SWE and AAE. Recall that speakers of these two dialects and children with and without DLD within these dialects differ in the frequency with which they produce overt forms and zero forms, and this difference may relate to their acceptability judgments of these forms.

Correlations Between Measures of Production and Acceptability

Table 5 reports the children's percentages of overt forms produced on the elicitation probes. Reported are the percentages for the 64 T/A items overall (i.e., items targeting past tense, verbal *–s*, BE present, and BE past) and for the items specifically targeting forms of *is* and verbal *–s*. As shown, dialectal differences were evident for all three production measures for the TD groups but not for the DLD groups. These findings mirror the results of the acceptability percentages in the grammaticality judgment task.

Spearman correlations were then completed to examine relations between the children's percentages of overt form productions and their acceptability percentages. As shown in Table 6, the results again depended on clinical status. For the TD groups, many statistically reliable correlations were found between the children's overt form productions and their judgments. Excluding the overt verbal *–s* forms, correlations between the TD groups' overt form productions and judgments of the overt forms were positive, and correlations between their overt form productions and judgments of the zero forms were negative. These results demonstrate a relation between the TD groups' productions and judgments.

The one exception to the TD groups' general pattern of results involved the overt verbal *–s* forms because no relation between production and grammaticality judgment was found for these forms. Recall that both the SWE and AAE TD groups accepted the overt verbal *–s* forms at high levels (90% and 87%, respectively), but they differed in their percentages of overt form productions (91% and 47%, respectively). These mixed findings—showing dialectal differences in production but not in judgment—explain the null correlation for these forms.

In stark contrast, hardly any correlations were observed for the DLD groups. Recall also that the DLD groups' A' scores and many of their acceptability percentages did not differ from chance. Thus, children with DLD had difficulty with grammatical judgment and with

Table 6. Correlations between acceptability percentages and overt form productions by clinical group (dialects combined).

Probe Items	Overt T/A forms	Overt *is*	Overt verbal *–s*	Zero *is*	Zero verbal *–s*	Zero progressive	Overt T/A misapplications
TD groups (n = 127)							
Production probe: All items	.19*	.24**	.03	–.44***	–.49***	–.32***	–.55***
Production probe: Overt *is*	.23**	.20*	.09	.27**	–.20*	–.20*	–.35***
Production probe: Overt verbal –s	.15	.22*	–.01	–.44***	–.46***	–.27**	–.52***
DLD groups (n = 52)							
Production probe: All structures	.23	.14	.34*	.05	.05	–.09	–.08
Production probe: Overt *is*	.20	.14	.22	–.04	.13	–.18	–.12
Production probe: Overt verbal –s	.29*	.15	.38**	.16	.03	.03	.07

Note. T/A = tense and agreement; TD = typically developing; DLD = developmental language disorder.
*p < 05. **p < .01. ***p < .001.

production, but these difficulties were largely not related to each other. Taken together, these findings underscore the extreme difficulty children with DLD have across SWE and AAE when asked to make judgments about sentences. In other words, although dialectal differences exist between SWE- and AAE-speaking children, these differences do not override the linguistic difficulties of children with DLD within AAE and SWE.

Discussion

GAE-speaking children with DLD perform more poorly on grammaticality judgment tasks than do GAE-speaking TD controls (Leonard et al., 2009; Miller et al., 2009; Noonan et al., 2014; Rice et al., 1999, 2009), and Vaughn et al. (2023) recently replicated this finding for AAE. In the current study, we further tested this finding by focusing on SWE. Compared to GAE, AAE and SWE include a greater variety of T/A form types, including not only dialect general overt forms (e.g., *walk/s*, *is walking*) that are shared across dialects but also dialect-specific zero forms (e.g., *walkØ, Ø walking)* that are not. AAE and SWE also differ from each other in the frequency with which overt (e.g., *walk/s*) and zero (e.g., *walkØ*) T/A forms are produced. Dialectal differences in form production may affect children's judgments of what is considered acceptable within their dialects and influence the types of structure- and form-specific judgments that differentiate children with and without DLD across dialects. The current study was designed to examine these possibilities.

The first research question focused on differences between the SWE DLD and TD groups. Results indicated that the SWE-speaking children with DLD struggled to judge various types of sentences, including those with T/A overt forms and zero forms. This finding was shown using A' scores and percentages of acceptability. All three SWE A' scores obtained by the DLD group were lower than those of the TD group, and none of the DLD A' scores were above chance or differed from each other. By comparison, the TD group's three A' scores were above chance, although their A' scores for the zero T/A forms were lower than the others. Recall that the zero T/A forms are grammatical in SWE as they are in AAE. Judging some of the zero T/A forms as grammatical would lower this A' score.

Clinical group differences were further documented using percentages of acceptability, and these metrics provided greater insight into the nature of the DLD weaknesses within SWE. Here, the DLD and TD groups' acceptability percentages differed on all structures and forms, with the DLD group accepting the overt T/A forms at levels above chance and all others at chance. By comparison, the TD group accepted the overt T/A forms above chance, judged the zero verbal *–s* forms at chance, and rejected the zero progressive forms and overt T/A misapplications. These results show the DLD group being less discriminating in their judgments than the TD group, even though they were able to accurately judge the overt T/A forms as acceptable for their dialect.

The second question focused on dialectal differences between the children's judgments. With the AAE and SWE dialect speakers combined, no dialectal differences were observed for either the TD or DLD groups' judgments of the overt forms, and the DLD groups also did not show a dialectal difference in their judgments of the other forms. The results differed for the TD groups because dialectal differences were revealed within their judgments of the zero *is* forms, zero progressive forms, and overt T/A misapplications. The SWE TD group's percentages of acceptability for each of these were also lower than the AAE TD group's, even though both groups rejected these forms. Most strikingly, however, were dialectal differences observed for the zero verbal *–s* forms, with the SWE TD group performing at chance and the AAE TD group accepting them at levels above chance. The AAE TD group's acceptance of the zero and overt verbal *–s* forms also did not differ from each other (zero: 72% vs. overt: 87% as reported by Vaughn et al., 2023), whereas they did for the SWE TD group (zero: 44% vs. overt: 90%).

The final research question focused on the relations between the children's grammaticality judgments and their production of T/A overt forms across the dialects. Relations were found, but they were primarily limited to the TD groups, and they were not observed for the overt verbal *–s* forms for this group. Interestingly, for all other structures and forms, correlations between the TD groups' judgments and productions were just as high when the children's production data included all the T/A items on the probes as when the production data were limited to the *is* and verbal *–s* items. This finding suggests that the TD groups' judgments of sentences with various types of overt forms and zero forms are related to their general production preferences across structures rather than tightly tied to their production preferences for specific structures.

Research and Clinical Implications

The findings further document the grammatical weaknesses of children with DLD across dialects of English, and they highlight the need for more cross-dialectal studies, especially for structures such as verbal *–s* that differ across dialects. Although a handful of child language studies have focused on verbal *–s* in AAE (e.g., Beyer & Hudson Kam, 2009; Green, 2019; Johnson, 2005;

Newkirk-Turner & Green, 2016; J. M. Terry et al., 2010, 2022; Van Hofwegen & Wolfram, 2010), SWE (e.g., Cleveland & Oetting, 2013; Hendricks & Adlof, 2020; Oetting et al., 2019, 2021) and GAE (e.g., Beyer & Hudson Kam, 2012; de Villiers & Johnson, 2007; Rice et al., 1999), they have not routinely examined the profile of childhood DLD in more than one dialect at a time. Hopefully, these early studies can be used to guide the next generation of larger, cross-dialectal efforts.

Cross-dialectal studies and the development of clinical resources generated from these studies need to be done now. The field of speech-language pathology has spent decades cautioning researchers and clinicians about mistakes we may make if we misinterpret a dialect difference (i.e., a linguistic behavior reflective of a dialect that differs from GAE) as a disorder (American Speech-Language-Hearing Association, 1983; Seymour et al., 1998). The need for researchers and clinicians to practice with caution will always be important. Indeed, had the GAE A' norms of the TEGI been utilized for the SWE- and AAE-speaking children studied here, far too many children in the TD groups (SWE: 33%–40%, AAE: 47%–63%) would have been incorrectly classified as impaired. However, the need to remain cautious should never serve as an excuse to not study or fully assess and treat the linguistic systems of children with DLD in dialects that differ from GAE. Unfortunately, concern over misinterpreting a dialect difference as a disorder within the field has contributed to clinical practices that do not allow a child's complete linguistic system to be evaluated. As an example, modified scoring systems within standardized language tests is a practice that was developed to avoid misinterpreting a dialect difference as a disorder. However, these scoring systems do not allow a clinician to assess (and learn about) a child's entire linguistic system; these scoring systems also do not lead to high levels of diagnostic accuracy (Hendricks & Adlof, 2017; Hendricks et al., 2023; Oetting et al., 2013, 2019, 2021; Oetting, Gregory, & Rivière, 2016; for other reasons the field lacks literature on AAE specifically, see Stockman, 2010). Modified scoring systems need to be significantly revised to help clinicians learn about the entire linguistic system of all children within the context of their dialect(s).

Many different types of cross-dialectal studies are needed, including those focused on TD groups, to better understand development and developmental milestones (e.g., Craig & Washington, 2002; Green, 2019; Newkirk-Turner & Green, 2016; Newkirk-Turner et al., 2015; Roy et al., 2013; Stockman et al., 2013), those focused on DLD and TD comparison groups to improve the accuracy at which children are diagnosed (e.g., Hendricks & Adlof, 2017, 2020; Hendricks et al., 2023; Maher et al., 2021; Overton et al., 2021), and those focused specifically on DLD groups to test the efficacy of interventions (e.g., Smith & Bellon-Harn, 2015). These types of studies need to be added to other, more established lines of research that examine linguistic biases within GAE-based assessments (e.g., Craig et al., 2004; Hendricks & Adlof, 2017); children's varying use of dialect-specific forms as they age (e.g., Horton-Ikard & Ellis Weismer, 2005; Ivy & Masterson, 2011); and the effects of dialect-specific form use on children's development of reading, writing, and other literacy-based tasks (e.g., Campbell et al., 2023; Fitton et al., 2021; N. P. Terry et al., 2012; Washington et al., 2018; for a meta-analysis, see also Gatlin & Wanzek, 2015).

The findings also speak to the importance of expanding the types of measures included within cross-dialectal research and clinical practice. Much of the research on childhood DLD across dialects has focused on measures of production. Findings from the current study indicate that measures of production cannot serve as a proxy for understanding a child's linguistic system. This is a key finding. Recall that correlations for the overt verbal *–s* forms within the TD groups did not pattern like the others, which highlights the need to study and assess these forms (and the zero verbal *–s* forms) with more rigor. In addition, for the other forms, the TD groups' judgments of the individual T/A forms were as related to their production of the T/A forms combined as they were to their production of the individual forms. Finally, these same correlations were not found for the DLD groups. Together, these findings highlight the importance of not making assumptions about children's dialects or their profile of childhood DLD within a dialect(s) without collecting multiple measures and using multiple types of tasks.

Expanding the research enterprise and clinical practice will require the development of new measures to better assess children's comprehension abilities and judgments. As an example, studies of dialect variation are beginning to combine electroencephalographic measures with behavioral measures (e.g., Garcia et al., 2022; J. M. Terry et al., 2022; Zaharchuk et al., 2021). If replicated across studies, this type of research may show promise for uncovering additional differences between children with and without DLD across dialects. To reiterate, it is very important to not expect children across dialects to always judge sentences in ways that match their productions. Other studies support this recommendation. Garcia et al. (2022) provided bidialectal AAE/GAE-speaking adults and monodialectal GAE-speaking adults accusative object versus nominative object pronouns (e.g., *The gentle doctor comforts them/they in the clinic*). Within AAE and GAE, accusative object pronouns are grammatical, whereas nominative object pronouns are not. Contrary to the authors' predictions, the bidialectal speakers judged the sentences with nominative object pronouns as more acceptable than the monodialectal speakers, and they took longer to judge

these sentences, even though they were ungrammatical in AAE and GAE. The authors hypothesized that this finding occurred because, compared to GAE, AAE offers a wider variety of pronoun options in other types of sentences (e.g., appositive pronouns—*my sister, she*, personal datives—*I got me a drink*, demonstrative them—*them cars*), and these other types of sentences impacted the AAE speaker's judgments of the nominative object pronouns.

Garcia et al.'s (2022) hypothesis, which we refer to as the associated structural variation account, may help explain why the AAE TD group studied here accepted the zero progressive forms and overt T/A misapplications at higher rates than the SWE TD group (even though both groups rejected these forms and the dialect difference for zero progressive form was not significant, $p = .06$). While the specific zero progressive forms targeted on the TEGI are ungrammatical in AAE and SWE, AAE allows to a greater degree than SWE reduced and zero progressive morphemes with the verb *go* (e.g., *gonna, gon,* and *I'ma*, as in *I'ma walk to the store*). Similarly, while the specific overt T/A misapplications (e.g., *he am*) targeted on the TEGI are ungrammatical in AAE and SWE, AAE allows overt forms of *is* and *was* to be produced with both singular and plural subjects (*he/they is*; *he/they was*) to a greater degree than SWE (Garrity & Oetting, 2010; Oetting et al., 2019, 2021). Thus, there is more associated structural variation within the progressive system and the BE system of AAE than SWE that could have influenced the AAE TD group's judgments of sentences.

Of course, the associated structural variation account does not fully explain the SWE and AAE DLD groups' chance performance with zero progressive forms and overt T/A misapplications or their lack of a correlation between their judgments and their productions. Recall that Rice et al. (1999) concluded that GAE-speaking children with DLD can judge sentences that are ungrammatical in their dialect if they do not involve zero forms of T/A. Thus, it is tempting to conclude that AAE- and SWE-speaking children with DLD show a different and perhaps more global language deficit than has been documented within GAE. However, at the age of the children studied here, the GAE-speaking DLD group studied by Rice et al. (1999) earned low *A′* scores for all sentence types tested, including those we refer to as overt T/A misapplications. In other studies, GAE-speaking children with DLD, aged 7–8 years, have also earned low *A′* scores for some grammatical errors that do not involve zero T/A forms (Wulfeck et al., 2004). Thus, these findings suggest that across dialects of English, the profile of DLD in young children includes weaknesses in the ability to make judgments about various types of grammar structures, despite the ability to demonstrate above chance performance on judgments of some forms that are grammatical in their dialect(s).

Finally, earlier we mentioned the need for a theoretical model that can explain both the typical variation that occurs between dialects and the atypical variation that occurs when children present with DLD within their dialect(s). Differences between dialects are ubiquitous to all language learners, including those with DLD. Nevertheless, for the grammaticality judgment task examined here, the linguistic weaknesses of the children with DLD overrode their dialectal differences. Dialogue, collaborative projects, and data sharing across fields are needed to develop a model(s) to account for these findings.

Limitations and Future Directions

Grammaticality judgments were measured using the TEGI, a test designed for speakers of GAE. As noted by Vaughn et al. (2023), the TEGI did not include enough items nor the appropriate types of items to evaluate the full repertoire of dialect-general and dialect-specific T/A forms found in SWE and AAE. The SWE- and AAE-inappropriate items on the TEGI were also limited in number and type. Given that the DLD groups studied here struggled to judge not only the T/A zero forms but also the dialect-inappropriate zero progressive forms and overt T/A misapplications, future grammaticality judgment studies need to include a greater number and a more diverse set of structures and forms.

Grammaticality judgments were measured using the options of *good* versus *not so good*. Pratt and Grinstead (2008) view this testing format as requiring unnecessary processing demands that can be reduced by asking children to select the best sentence from a field of two (one grammatical and the other ungrammatical). The binary format of the task also did not allow the children to provide gradient responses. Other studies involving adults do this. For example, Blanchette (2017) asked participants to rate the naturalness of various types of sentences with negative concord (e.g., *didn't eat nothing*) using a 7-point scale. Zaharchuk et al. (2021) asked participants to indicate the intelligibility, acceptability, and familiarity of various types of sentences with double modals (e.g., *might could*) using 4- and 5-point scales. These types of gradient measures could be simplified and developed for children to learn more about their grammar preferences as a function of their dialect and clinical status. Gradient measures may also help reveal changes in children's judgments as they age.

Conclusions

As found for AAE, the SWE DLD group was less able than the TD group to judge various types of sentences

including those with T/A structures as measured by A' scores and percentages of acceptability. In addition, with the dialects combined, only the judgments of the TD groups showed dialect effects, particularly for the verbal *–s* forms that differ in their frequency within SWE and AAE. This finding is interesting and needs to be further tested because it suggests that dialect differences between SWE- and AAE-speaking children do not always override the linguistic difficulties of children with DLD within these dialects. More cross-dialectal research is needed to better understand the grammatical weaknesses of childhood DLD, especially for structures like verbal *–s* that are expressed differently across dialects of English.

Author Contributions

Janna B. Oetting: Conceptualization (Co-Lead), Data curation (Lead), Methodology (Co-Lead), Project administration (Co-Lead), Funding acquisition (Co-Lead), Writing – original draft (Lead), Writing – review & editing (Lead). **Janet L. McDonald:** Conceptualization (Co-Lead), Data curation (Supporting), Formal analysis (Lead), Project administration (Co-Lead), Funding acquisition (Co-Lead), Methodology (Co-Lead), Writing – original draft (Supporting), Writing – review & editing (Supporting). **Lori E. Vaughn:** Methodology (Supporting), Formal analysis (Supporting), Validation (Lead), Writing – original draft (Supporting), Writing – review & editing (Supporting).

Data Availability Statement Data

The grammaticality judgment data file may be requested from the authors.

Acknowledgments

The research was funded by National Institute on Deafness and Other Communication Disorders Grant RO1DC009811 awarded to Janna B. Oetting, Janet L. McDonald, and Michael Hegarty. Appreciation is extended to Jessica Berry, Kyomi Gregory-Martin, Ryan Lee-James, Andy Rivière, Christy Seidel, Tina Villa, and other students who collected the data. The authors would also like to thank the teachers, families, and children who participated in the study.

References

American Speech-Language-Hearing Association. (1983). *Social dialects* [Position statement]. https://www.asha.org/policy/PS1983-00115

Beyer, T., & Hudson Kam, C. (2009). Some cues are stronger than others: The (non)interpretation of 3rd person present *–s* as a tense marker by 6- and 7-year-olds. *First Language, 29*(2), 208–227. https://doi.org/10.1177/0142723708101678

Beyer, T., & Hudson Kam, C. (2012). First and second graders' interpretation of standard American English morphology across varieties of English. *First Language, 32*(3), 365–384. https://doi.org/10.1177/0142723711427618

Bishop, D. 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 and Allied Disciplines, 58*(10), 1068–1080. https://doi.org/10.1111/jcpp.12721

Blanchette, F. (2017). Micro-syntactic variation in American English negative concord. *Glossa: A Journal of General Linguistics, 2*(1), Article 65. https://doi.org/10.5334/gigl.188

Campbell, D., Wood, C., & Hall-Mills, S. (2023). An examination of 3rd and 5th grade students' use of dialect specific forms during a written editing task. *Journal of Communication Disorders, 102,* Article 106303. https://doi.org/10.1016/j.jcomdis.2023.106303

Cleveland, L. H., & Oetting, J. B. (2013). Children's marking of verbal –s by nonmainstream English dialect and clinical status. *American Journal of Speech-Language Pathology, 22*(4), 604–614. https://doi.org/10.1044/1058-0360(2013/12-0122)

Craig, H. K., Thompson, C. A., Washington, J. A., & Potter, S. L. (2004). Performance of elementary-grade African American students on the Gray Oral Reading Tests. *Language, Speech, and Hearing Services in Schools, 35*(2), 141–154. https://doi.org/10.1044/0161-1461(2004/015)

Craig, H. K., & Washington, J. A. (2002). Oral language expectations for African American preschoolers and kindergartners. *American Journal of Speech-Language Pathology, 11*(1), 59–70. https://doi.org/10.1044/1058-0360(2002/007)

de Villiers, J. G., & Johnson, V. E. (2007). The information in third-person/s: Acquisition across dialects of American English. *Journal of Child Language, 34*(1), 133–158. https://doi.org/10.1017/s0305000906007768

Dunn, L. M., & Dunn, D. M. (2007). *Peabody Picture Vocabulary Test–Fourth Edition*. PsychCorp.

Ehrler, D. J., & McGhee, R. L. (2008). *Primary Test of Nonverbal Intelligence*. Pro-Ed.

Fitton, L., Johnson, L., Wood, C., Schatschneider, C., & Hart, S. A. (2021). Language variation in the writing of African American students: Factors predicting reading achievement. *American Journal of Speech-Language Pathology, 30*(6), 2653–2667. https://doi.org/10.1044/2021_AJSLP-20-00263

Garcia, F. M., Shen, G., Avery, T., Green, H. L., Godoy, P., Khamis-Dakwar, R., & Froud, K. (2022). Bidialectal and monodialectal differences in morphosyntactic processing of AAE and MAE: Evidence from ERPs and acceptability judgments. *Journal of Communication Disorders, 100*(5), Article 106267. https://doi.org/10.1016/j.jcomdis.2022.106267

Garrity, A. W., & Oetting, J. B. (2010). Auxiliary BE production by African American English–speaking children with and without specific language impairment. *Journal of Speech, Language, and Hearing Research, 53*(5), 1307–1320. https://doi.org/10.1044/1092-4388(2010/09-0016)

Gatlin, B., & Wanzek, J. (2015). Relations among children's use of dialect and literacy skills: A meta-analysis. *Journal of Speech, Language, and Hearing Research, 58*(4), 1306–1318. https://doi.org/10.1044/2015_JSLHR-L-14-0311

Goldman, R., & Fristoe, M. (2000). *Goldman-Fristoe Test of Articulation–Second Edition*. AGS.

Green, L. J. (2019). All zeros are not equal in African American English. In D. W. Lightfoot & J. Havenhill (Eds.), *Variable properties in language: Their nature and acquisition* (pp. 183–194). Georgetown University Press. https://doi.org/10.2307/j.ctvfxv99p.19

Hendricks, A. E., & Adlof, S. M. (2017). Language assessment with children who speak nonmainstream dialects: Examining the effects of scoring modifications in norm-referenced assessment. *Language, Speech, and Hearing Services in Schools, 48*(3), 168–182. https://doi.org/10.1044/2017_LSHSS-16-0060

Hendricks, A. E., & Adlof, S. M. (2020). Production of morphosyntax within and across different dialects of American English. *Journal of Speech, Language, and Hearing Research, 63*(7), 2322–2333. https://doi.org/10.1044/2020_JSLHR-19-00244

Hendricks, A. E., Jerard, J., & Guo, L. (2023). Evaluating different scoring systems for a picture description task among preschool children who speak African American English. *Language, Speech, and Hearing Services in Schools, 54*(1), 198–211. https://doi.org/10.1044/2022_LSHSS-21-00189

Horton-Ikard, R., & Ellis Weismer, S. (2005). Distinguishing African-American English from developmental errors in the language production of toddlers. *Applied Psycholinguistics, 26*(4), 597–620. https://doi.org/10.1017/S0142716405050320

Ivy, L. J., & Masterson, J. J. (2011). A comparison of oral and written English styles in African American students at different stages of writing development. *Language, Speech, and Hearing Services in Schools, 42*(1), 31–40. https://doi.org/10.1044/0161-1461(2010/09-0069)

Johnson, V. E. (2005). Comprehension of third person singular /s/ in AAE-speaking children. *Language, Speech, and Hearing Services in Schools, 36*(2), 116–124. https://doi.org/10.1044/0161-1461(2005/011)

Leonard, L. B. (2017). *Children with specific language impairment*. MIT Press.

Leonard, L. B., Miller, C. A., & Finneran, D. A. (2009). Grammatical morpheme effects on sentence processing by school aged adolescents with specific language impairment. *Language and Cognitive Processes, 24*(3), 450–478. https://doi.org/10.1080/01690960802229649

Maher, Z. K., Erskine, M. E., Byrd, A. S., Harring, J. R., & Edwards, J. R. (2021). African American English and early literacy: A comparison of approaches to quantifying nonmainstream dialect use. *Language, Speech, and Hearing Services in Schools, 52*(1), 118–130. https://doi.org/10.1044/2020_LSHSS-19-00115

Miller, C. A., Leonard, L. B., & Finneran, D. (2009). Grammaticality judgements in adolescents with and without language impairment. *International Journal of Language & Communication Disorders, 43*(3), 346–360. https://doi.org/10.1080/13682820701546813

Newkirk-Turner, B. R., & Green, L. (2016). Third person singular *–s* and event marking in child African American English. *Linguistic Variation, 16*(1), 103–130. https://doi.org/10.1075/lv.16.1.05new

Newkirk-Turner, B. R., Horton, R. A., & Stockman, I. J. (2015). Language acquisition in African American English prior to age four. In S. Lanehart (Ed.), *African American language* (pp. 439–544). Oxford University Press.

Noonan, N. B., Redmond, S. M., & Archibald, L. M. (2014). Contributions of children's linguistic and working memory proficiencies to their judgments of grammaticality. *Journal of Speech, Language, and Hearing Research, 57*(3), 979–989. https://doi.org/10.1044/2014_JSLHR-L-12-0225

Oetting, J. B. (2019). Variability within varieties of language: Profiles of typicality and impairment. In T. Ionin & M. Rispoli (Eds.), *Selected proceedings of the 7th Generative Approaches to Language Acquisition—North America Conference* (pp. 59–82). John Benjamins.

Oetting, J. B., Berry, J. R., Gregory, K. D., Rivière, A. M., & McDonald, J. (2019). Specific language impairment in African American English and Southern White English: Measures of tense and agreement with dialect-informed probes and strategic scoring. *Journal of Speech, Language, and Hearing Research, 62*(9), 3443–3461. https://doi.org/10.1044/2019_JSLHR-L-19-0089

Oetting, J. B., Berry, J. R., & Gregory-Martin, K. (2022). Use of linguistic theory to inform the assessment and treatment of children with developmental language disorder within African American English. In N. Gurevich & C. Grindrod (Eds.), *Clinical applications of linguistics to speech-language pathology: A guide for clinicians* (pp. 72–90). Routledge. https://doi.org/10.4324/9781003045519-5

Oetting, J. B., & Garrity, A. W. (2006). Variation within dialects: A case of Cajun/Creole influence within child SAAE and SWE. *Journal of Speech, Language, and Hearing Research, 49*(1), 16–26. https://doi.org/10.1044/1092-4388(2006/002)

Oetting, J. B., Gregory, K. D., & Rivière, A. M. (2016). Changing how speech-language pathologists think and talk about dialect variation. *Perspectives of the ASHA Special Interest Groups, 1,* 28–37. https://doi.org/10.1044/persp1.SIG16.28

Oetting, J. B., Lee, R., & Porter, K. (2013). Evaluating the grammars of children who speak nonmainstream dialects of English. *Topics in Language Disorders, 33*(2), 140–151. https://doi.org/10.1097/TLD.0b013e31828f509f

Oetting, J. B., & McDonald, J. L. (2001). Nonmainstream dialect use and specific language impairment. *Journal of Speech, Language, and Hearing Research, 44*(1), 207–223. https://doi.org/10.1044/1092-4388(2001/018)

Oetting, J. B., & McDonald, J. L. (2002). Methods for characterizing participants' nonmainstream dialect use in child language research. *Journal of Speech, Language, and Hearing Research, 45*(3), 505–518. https://doi.org/10.1044/1092-4388(2002/040)

Oetting, J. B., McDonald, J. L., Seidel, C. M., & Hegarty, M. (2016). Sentence recall by children with SLI across two nonmainstream dialects of English. *Journal of Speech, Language, and Hearing Research, 59*(1), 183–194. https://doi.org/10.1044/2015_JSLHR-L-15-0036

Oetting, J. B., Rivière, A. M., Berry, J. R., Gregory, K. D., Villa, T. M., & McDonald, J. (2021). Marking of tense and agreement in language samples by children with and without specific language impairment in African American English and Southern White English: Evaluation of scoring approaches and cut scores across structures. *Journal of Speech, Language, and Hearing Research, 64*(2), 491–509. https://doi.org/10.1044/2020_JSLHR-20-00243

Overton, C., Taylor, B., Person, B. Z., & Ratner, N. B. (2021). Using free computer-assisted language sample analysis to evaluate and set treatment goals for children who speak African American English. *Language, Speech, and Hearing Services in Schools, 52*(1), 31–50. https://doi.org/10.1044/2020_LSHSS-19-00107

Pratt, A., & Grinstead, J. (2008). Receptive measures of the optional infinitive stage in child Spanish. In J. Bruhn de Garavito, & E. Valenzuela (Eds.), *Selected proceedings of the 10th Hispanic Linguistics Symposium* (pp. 120–133). Cascadilla Press.

Pruitt, S., Garrity, A. W., & Oetting, J. B. (2010). Family history of speech and language impairment in African American children. *Topics in Language Disorders, 30*(2), 154–164. https://doi.org/10.1097/TLD.0b013e3181e03f9

Rice, M. L. (2003). A unified model of specific and general language delay: Grammatical tense as a clinical marker of unexpected variation. In Y. Levy & J. Schaeffer (Eds.), *Language competence across populations: Toward a definition of specific language impairment* (pp. 63–95). Erlbaum.

Rice, M. L., Hoffman, L., & Wexler, K. (2009). Judgments of omitted BE and DO in questions as extended finiteness clinical markers of specific language impairment (SLI) to 15 years: A study of growth and asymptote. *Journal of Speech, Language, and Hearing Research, 52*(6), 1417–1433. https://doi.org/10.1044/1092-4388(2009/08-0171)

Rice, M. L., & Wexler, K. (1996). Toward tense as a clinical marker of specific language impairment in English-speaking children. *Journal of Speech and Hearing Research, 39*(6), 1239–1257. https://doi.org/10.1044/jshr.3906.1239

Rice, M. L., & Wexler, K. (2001). *Rice/Wexler Test of Early Grammatical Impairment*. The Psychological Corporation.

Rice, M. L., Wexler, K., & Cleave, P. (1995). Specific language impairment as a period of extended optional infinitive. *Journal of Speech and Hearing Research, 38*(4), 850–863. https://doi.org/10.1044/jshr.3804.850

Rice, M. L., Wexler, K., & Hershberger, S. (1998). Tense over time: The longitudinal course of tense acquisition in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 41*(6), 1412–1431. https://doi.org/10.1044/jslhr.4106.1412

Rice, M. L., Wexler, K., & Redmond, M. S. (1999). Grammaticality judgments of an extended optional infinitive grammar: Evidence from English-speaking children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 42*(4), 943–961. https://doi.org/10.1044/jslhr.4204.943

Rivière, A. M., Oetting, J. B., & Roy, J. (2018). Effects of specific language impairment on a contrastive dialect structure: The case of infinitival TO across various nonmainstream dialects of English. *Journal of Speech, Language, and Hearing Research, 61*(8), 1989–2001. https://doi.org/10.1044/2018_JSLHR-L-17-0209

Roy, J., Oetting, J. B., & Moland, C. (2013). Linguistic constraints on children's overt marking of BE by dialect and age. *Journal of Speech, Language, and Hearing Research, 56*(3), 933–944. https://doi.org/10.1044/1092-4388(2012/12-0099)

Seymour, H. N., Bland-Stewart, L., & Green, L. J. (1998). Difference versus deficit in child African American English. *Language, Speech, and Hearing Services in Schools, 29*(2), 96–108. https://doi.org/10.1044/0161-1461.2902.96

Seymour, H. N., Roeper, T., & de Villiers, J. (2003). *Diagnostic Evaluation of Language Variation–Screening Test*. Ventris Learning.

Seymour, H. N., Roeper, T., & de Villiers, J. (2005). *Diagnostic Evaluation of Language Variation–Norm-Referenced Test*. Ventris Learning.

Shin, N. (2022). Structured variation in child heritage speakers' grammars. *Compass, 16*(12), Article e12480. https://doi.org/10.1111/lnc3.12480

Shin, N., & Miller, K. (2022). Children's acquisition of morphosyntactic variation. *Language Learning and Development, 18*(2), 125–150. https://doi.org/10.1080/15475441.2021.1941031

Smith, S., & Bellon-Harn, M. L. (2015). Rates of auxiliary is and are in African American English–speaking children with specific language impairment following language treatment. *Clinical Linguistics & Phonetics, 29*(2), 131–149. https://doi.org/10.3109/02699206.2014.966394

Stockman, I. J. (2010). A review of developmental and applied language research on African American children: From a deficit to difference perspective on dialect differences. *Language, Speech, and Hearing Services in Schools, 41*(1), 23–38. https://doi.org/10.1044/0161-1461(2009/08-0086)

Stockman, I. J., Guillory, B., Seibert, M., & Boult, J. (2013). Toward validation of a minimal competence core of morphosyntax for African American children. *American Journal of Speech-Language Pathology, 22*(1), 40–56. https://doi.org/10.1044/1058-0360(2012/11-0124)

Terry, J. M., Jackson, S. C., Evangelou, E., & Smith, R. L. (2010). Expressive and receptive language effects of African American English on a sentence imitation task. *Topics in Language Disorders, 30*(2), 119–134. https://doi.org/10.1097/TLD.0b013e3181e04148

Terry, J. M., Thomas, E. R., Jackson, S. C., & Hirotoni, M. (2022). African American English speaking 2nd graders, verbal *–s*, and educational achievement: Event related potential and math study findings. *PLOS ONE, 17*(10), Article e0273926. https://doi.org/10.1371/journal.pone.0273926

Terry, N. P., Connor, C. M., Petsher, Y., & Conlin, C. R. (2012). Dialect variation and reading: Is change in nonmainstream American English use related to reading achievement in first and second grades? *Journal of Speech, Language, and Hearing Research, 55*(1), 55–69. https://doi.org/10.1044/1092-4388(2011/09-0257)

Van Hofwegen, J., & Wolfram, W. (2010). Coming of age in African American English: A longitudinal study. *Journal of Sociolinguistics, 14*(4), 427–455. https://doi.org/10.1111/j.1467–9841.2010.00452.x

Vaughn, L., Oetting, J. B., & McDonald, J. L. (2023). Grammaticality judgments of tense and agreement by child speakers of African American English: Effects of clinical status, surface form, and grammatical structure. *Journal of Speech, Language, and Hearing Research, 66*(5), 1755–1770. https://doi.org/10.1044/2023_JSLHR-22-00431

Washington, J. A., Branum-Martin, L., Sun, C., & Lee-James, R. (2018). The impact of dialect density on the growth of language and reading in African American children. *Language, Speech, and Hearing Services in Schools, 49*(2), 232–247. https://doi.org/10.1044/2018_LSHSS-17-0063

Wulfeck, B., Bates, E., Krupa-Kwiatkowski, M., & Saltzman, D. (2004). Grammaticality sensitivity in children with early focal brain injury and children with specific language impairment. *Brain and Language, 88*(2), 215–228. https://doi.org/10.1016/S0093-934X(03)00100-7

Zaharchuk, H. A., Shevlin, A., & van Hell, J. G. (2021). Are our brains more prescriptive than our mouths? Experience with dialectal variation in syntax differentially impacts ERPs and behavior. *Brain and Language, 218,* Article 104949. https://doi.org/10.1016/j.bandl.2021.104949

Copyright of Journal of Speech, Language & Hearing Research is the property of American Speech-Language-Hearing Association and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.