Review Article

Past Tense Production in Children With and Without Specific Language Impairment Across Germanic Languages: A Meta-Analysis

Windi C. Krok^a and Laurence B. Leonard^a

Purpose: This study examined the extent to which children with specific language impairment (SLI) across Germanic languages differ from their typically developing (TD) peers in the use of past tense morphology.

Method: A systematic literature search identified empirical studies examining regular and/or irregular past tense production by English and non-English Germanic-speaking children with SLI and their TD peers. Data from qualifying studies were extracted and converted to Hedges's *g* effect sizes.

Results: Seventeen English and 8 non-English Germanic studies met inclusionary criteria. Comparing children with SLI and their TD age-matched (TDA) peers resulted in large

combined effect sizes for English and non-English Germanic regular and irregular past tense production. Comparisons between children with SLI and their TD younger (TDY) peers also revealed large combined effect sizes for English and non-English Germanic regular past tense production. Effect sizes for studies comparing SLI and TDY irregular past tense production were large for non-English Germanic-speaking children and moderate for English-speaking children.

Conclusions: Results suggest that children with SLI across Germanic languages do indeed have more difficulty marking verbs for past tense than TDA and TDY peers. The findings suggest that the potential value of past tense production as a clinical marker of SLI may well extend beyond English.

Children with specific language impairment (SLI) demonstrate impaired language abilities in the absence of hearing, cognitive, neurological, and emotional impairments (Leonard, 2014). Many children with SLI have disproportionately more difficulty with grammar than other areas of language. For children with SLI acquiring English, tense and subject–verb agreement morphemes, such as copula and auxiliary *be* forms, auxiliary *do* forms, present third person singular *-s*, and past tense *-ed*, are particularly difficult (Rice, Wexler, & Hershberger, 1998). These tense/agreement morphemes emerge later in the speech of children with SLI than in typically developing (TD) children, and, once they emerge, they are used only inconsistently for a protracted period (Rice, Wexler, & Cleave, 1995). A weakness in the production of tense/agreement morphemes continues through the preschool years and

beyond (Rice et al., 1998) and has been proposed as a clinical marker of this disorder (Bedore & Leonard, 1998; Conti-Ramsden, Botting, & Faragher, 2001; Gladfelter & Leonard, 2013; Rice & Wexler, 1996).

English is not the only language in which special difficulties with tense/agreement morphology are seen in children with SLI. In other Germanic languages, children with SLI have considerable difficulty with these morphemes. Investigators have reported such weaknesses in Dutch (de Jong, 1999; Spoelman & Bol, 2012; Verhoeven, Steenge, & van Balkom, 2011; Wexler, Schaeffer, & Bol, 2004), Danish (Lum & Bleses, 2012; Vang Christensen & Hansson, 2012), German (Clahsen, Bartke, & Göllner, 1997; Rice, Noll, & Grimm, 1997; Rothweiler & Clahsen, 1994), Icelandic (Thordardottir, 2008), Norwegian (Simonsen & Bjerkan, 1998), and Swedish (Hansson, 1997; Hansson & Leonard, 2003; Hansson & Nettelbladt, 1995; Leonard, Hansson, Nettelbladt, & Deevy, 2004).

A different profile is seen in other types of languages. For example, in Romance languages that permit null subjects, such as Italian and Spanish, tense/agreement verb inflections do not appear to be any more difficult for children with SLI than other aspects of grammar (Bedore &

^aPurdue University, West Lafayette, IN

Correspondence to Windi C. Krok: wkrok@purdue.edu

Editor: Rhea Paul

Associate Editor: Shelley Gray

Received December 11, 2014

Revision received March 24, 2015

Accepted May 11, 2015

DOI: 10.1044/2015_JSLHR-L-14-0348

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

Leonard, 2001; Bedore & Leonard, 2005; Jacobsen & Schwartz, 2002; Leonard, Bortolini, Caselli, McGregor, & Sabbadini, 1992). On the other hand, function words that express tense/agreement (e.g., copula and auxiliary forms) are often quite problematic for these children (Leonard & Bortolini, 1998; Leonard, Sabbadini, Leonard, & Volterra, 1987; Paradis & Crago, 2001).

In this review article, we focus on one particular tense/agreement morpheme type—past tense, in both its regular (e.g., *played*) and irregular (e.g., *threw*) manifestations. There are numerous accounts of grammatical deficits in SLI that are applicable to these children's weaknesses with past tense. The assumptions underlying these accounts vary widely. At one end of the continuum is the assumption that children with SLI have difficulty acquiring implicit grammatical rules and must learn verbs inflected for regular past tense (e.g., *walked*) in the same item-by-item manner in which irregular past forms are learned (Gopnik & Crago, 1991).

According to the accounts offered by van der Lely (2004, 2005), the difficulty with regular past tense rests in the fact that it, like certain other linguistic operations, requires complex computations. In the linguistic framework adopted by van der Lely (2004) (that of Chomsky, 1995), a past tense verb in English must move from the syntactic position of verb to the position of inflection (INFL), or to have checking occur at INFL. Faced with this computational challenge, children with SLI resort to rote learning—a process that results in their learning (by rote) past tense forms of high frequency of occurrence more readily than past tense forms of lower frequency of occurrence.

Whereas van der Lely (2004) viewed tense problems in SLI as part of a wider (computational) problem within language, Ullman and Pierpoint (2005) regarded these difficulties as part of an even broader deficit. According to these researchers, the problems reflect a problem within the procedural memory system. As a result, weaknesses in children with SLI are seen in a wide range of activities that require the learning of sequential material, in nonlinguistic as well as linguistic domains.

Rice and colleagues (Rice & Wexler, 1996; Rice et al., 1998) view the past tense difficulties in a different way. According to these investigators, children with SLI possess the linguistic knowledge of past tense. However, these children are very late in acquiring a biologically based principle that makes tense obligatory in all main clauses. As a result, these children go through an extended period of inconsistent use, alternating between appropriate use of past tense and inappropriate production of nonfinite verbs in the same contexts (e.g., *Yesterday Jackie watched a movie*, *Yesterday Jackie watch a movie*). Although slight variations of this approach have appeared in the literature (Wexler, 2003; Wexler, Schütze, & Rice, 1998), its major assumptions are captured in its original name, the Extended Optional Infinitive account.

Other accounts emphasize processing limitations as the likely source of weaknesses with past tense. Leonard, Eyer, Bedore, and Grela (1997) proposed that the processing limitations of these children significantly hamper the

joint operations of perceiving inflections and hypothesizing their grammatical functions, especially when these inflections are brief in duration. Joanisse and Seidenberg (1998) attributed past tense weaknesses in SLI to these children's difficulties with phonologically segmenting words, resulting in the children learning items such as *walked* without recognizing their stem + past tense composition.

Yet another type of account is seen in the work of Leonard, Deevy, and their colleagues (Leonard & Deevy, 2011; Leonard, Fey, Deevy, & Bredin-Oja, 2015). In this account, children with SLI are assumed to alternate between utterances such as *The girl played tennis* and *The girl play tennis* because, along with hearing past tense verbs in simple declarative sentences, they also hear nonfinite verb forms (such as *play*) in structures such as *We saw the girl play tennis* and *Did the girl play tennis?* Not understanding that the nonfinite clauses are permissible only when preceded by finite information (e.g., *saw*, *did* in these examples), these children inappropriately extract the nonfinite clauses and use them as a basis for generating new utterances.

Each of these hypotheses can be evaluated to some degree by examining both regular past and irregular past tense verbs, not only in English but also in non-English Germanic languages. For example, irregular past tense forms should be less severely impaired than regular past tense forms if the source of difficulty were computational complexity, or learning implicit rules, or processing brief grammatical morphemes. Some of the differences between English and other Germanic languages can also serve to evaluate the alternative accounts of past tense difficulty. For example, some of the regular past tense inflections in non-English Germanic languages (e.g., Dutch, Swedish), are syllabic. Therefore, according to accounts that emphasize the importance of duration or production complexity, children with SLI in these languages should be less adversely affected than English-speaking children with SLI, who must deal with the frequent consonantal allomorphs of regular past tense, {t} and {d}.

All of these accounts assume that children with SLI make less use of irregular past tense and/or regular past tense than children with typical development who are matched for age (hereafter, TDA children) as well as younger children exhibiting typical development (hereafter, TDY children) who are matched on a measure such as mean length of utterance (MLU). The accounts differ in whether the same degree of difficulty will be seen across Germanic languages, and whether the problem will be confined to regular past tense.

The accumulating evidence provides preliminary support for the view that children with SLI do lag behind both TDA and TDY children in more than one Germanic language and that weaknesses may extend to irregular past tense. The purpose of this study is to put these impressions to a more rigorous test. By applying meta-analyses to the published data, we attempt to arrive at a more definitive interpretation of the status of past tense in children with SLI who speak these languages, and, in so doing, obtain a broader basis for evaluating some of the alternative proposals that have been offered to explain these deficits.

Method

Literature Search Procedures

Systematic literature searches for empirical articles were conducted between March and August 2014. The following four searches were conducted on EBSCO Host using the Academic Search Premier, Child Development and Adolescent Studies, Cumulative Index to Nursing and Allied Health Literature (CINAHL), Education Resources Information Center (ERIC), Medline, and PsycINFO databases: (1) “SLI” OR “specific language impairment” OR “specific-language impairment” OR “language impairment” OR “language delay” OR “developmental dysphasia” OR “primary language impairment” OR “developmental language delay” OR “developmental language impairment” or “language impair*”; (2) “past tense” OR “past” OR “-ed”; (3) “English”; and (4) “Danish” OR “Dutch” OR “German” OR “Icelandic” OR “Norwegian” OR “Swedish.” Combining searches 1, 2, and 3 yielded 322 English studies; combining searches 1, 2, and 4 yielded 18 non-English Germanic studies.

Inclusionary Criteria

To be included in the analysis, studies were required to meet the following inclusionary criteria: (a) written in English; (b) reported nonoverlapping empirical research data; (c) examined production of regular and/or irregular past tense verbs in monolingual Danish-, Dutch-, English-, German-, Icelandic-, Norwegian-, or Swedish-speaking children; (d) compared children with SLI with at least one group of children with typical development who were monolingual speakers of the same language; and (e) reported sufficient statistical information to calculate effect sizes for regular and/or irregular past tense production. Included studies were required to provide data for regular and/or irregular past tense separately rather than a single combined past tense score. Candidate studies were not restricted to those focusing on mainstream dialects of the designated languages. However, none of the studies that met our selection criteria specified the inclusion of children who spoke nonmainstream dialects.

Included Studies

A first-pass inspection eliminated duplicate and non-empirical research citations, resulting in 165 English and 15 non-English Germanic studies. When participant groups overlapped across studies, the study with the largest participant group was selected for inclusion. A second-pass inspection eliminated an additional 149 English and 10 non-English Germanic studies for not meeting the inclusionary criteria. One additional English and three additional non-English Germanic studies discovered in the reference sections of previously identified studies were added to the analysis. The final meta-analysis included 17 English studies and eight non-English Germanic studies (see Figures 1 and 2).

Data Selection

The identified studies varied with respect to elicitation methods and scoring of past tense verbs (see Tables 1 and 2). The included studies used elicitation probes ($n = 18$ studies), conversational language samples ($n = 4$ studies), a combination of probes and conversation ($n = 1$ study), and narrative tasks ($n = 2$ studies) to elicit past tense production from the participants. Each of these methods was accepted for inclusion in the meta-analysis. When two or more methods were reported, elicitation scores were selected unless a clear description of the elicitation method was not available or sufficient data for calculating effect sizes were not reported. Conversational samples consisted primarily of interactions between the child and a researcher. One study (Hansson & Nettelbladt, 1995) combined samples of the child interacting with a researcher, a parent, and a peer. Leonard et al. (2007) reported separate results for participants who did and did not achieve 100% accuracy on the production tasks. The results for participants who did not achieve 100% were chosen for this meta-analysis because it was the most complete data set. Further inspection of the data revealed that this decision did not affect overall effect size calculations. Only measurements from the first time point were selected from longitudinal studies.

Many studies reported multiple types of past tense production measures. In these cases, only overall production of regular and/or irregular past tense in real words was included in the meta-analysis. When results were divided into categories (e.g., high and low frequency regular past tense verbs) and an overall regular or irregular score was not reported, data from the category with the most conservative effect size were selected for inclusion in the meta-analysis. The selected regular past tense data included measures for the percentage or number of correct or omitted regular past tense markers. The regular past tense studies included 15 comparisons between SLI and TDA groups (10 English, five non-English) and 17 comparisons between SLI and TDY groups (12 English, five non-English). The selected irregular past tense data included the percentage or number of irregular past tense verbs marked or unmarked for tense. It is important to point out that *marked for tense* was defined here as correct irregular past form (e.g., *threw*) or overregularization (e.g., *throwed*). This definition placed a premium on the marking of tense independent of children's familiarity with the particular form that past tense takes in a particular verb. This definition resulted in the inclusion of five studies (four English, one non-English) for the SLI–TDA comparison, and six studies (four English, two non-English) for the SLI–TDY comparison. The remaining studies excluded overregularizations or treated them as errors and were, therefore, excluded from the meta-analysis.

Participants

A total of 494 children with SLI, 405 age-matched peers, and 344 younger peers participated across 25 studies

Figure 1. English study eligibility flowchart. CINAHL = Cumulative Index to Nursing and Allied Health Literature; ERIC = Education Resources Information Center; SLI = specific language impairment; TD = typically developing.

(see Tables 1 and 2). The participants ranged in age from 2;1 [years;months] to 17;8 (SLI: $M = 6;11$, $SD = 2;6$, TDA: $M = 6;4$, $SD = 1;10$; TDY: $M = 4;3$, $SD = 1;11$). Participant gender was reported for 86% of the SLI groups and 76% of the TD groups. Where reported, there were more boys than girls represented in each participant group (SLI: $M = 14$ boys, 8 girls; TDA: $M = 14$ boys, 11 girls; TDY: $M = 9$ boys, 8 girls).

SLI

In 24/25 (96%) studies, participants with SLI were identified on the basis of inclusionary and exclusionary criteria typical of an SLI diagnosis. Although children in 21/25 (84%) studies had a previous diagnosis of SLI upon recruitment, in all 25 studies the researchers conducted testing to verify language status. Children with SLI demonstrated MLUs and/or standardized language test scores

falling ≥ 1 to 1.5 SD s below the mean. They also demonstrated normal hearing and nonverbal intelligence, and a negative history of neurological and emotional delays. In the study that did not specify the inclusionary and exclusionary criteria used (Verhoeven et al., 2011), all children were nevertheless evaluated and judged by a multidisciplinary team of professionals to demonstrate SLI.

TDA

The TDA participants demonstrated MLUs and/or standardized test scores falling within the normal range. They also demonstrated normal hearing and nonverbal intelligence, and a negative history of neurological and emotional delays. One study (Verhoeven et al., 2011) did not describe the testing procedures used for identifying TD children. In all studies, the TDA participants were selected to resemble the children in the SLI group in age range. However,

Figure 2. Non-English Germanic study eligibility flowchart. CINAHL = Cumulative Index to Nursing and Allied Health Literature; ERIC = Education Resources Information Center; SLI = specific language impairment; TD = typically developing.

when averaged across studies, the children in the SLI groups were approximately 7 months older than those in the TDA groups.

TDY

The TDY participants also scored within the normal range according to their MLUs and/or standardized test scores. Like the TDA participants, these children also demonstrated normal hearing and nonverbal intelligence, along with a negative history of neurological and emotional delays. The TDY participants were at least 1 year younger than the SLI groups. In 14 of the 17 studies, the TDY participants were selected to resemble their SLI counterparts according to MLU ($n = 9$ studies), general morphosyntactic abilities ($n = 3$ studies), *language age* ($n = 1$ study), or

productive vocabulary score ($n = 1$ study). For the three remaining studies, the children were selected simply on the basis of being younger and scoring within the normal range on standardized test scores. When multiple TDY groups were included ($n = 1$ study), the group most closely resembling the SLI group on grammar rather than vocabulary was selected for the meta-analysis. This group was the youngest of the comparison groups and thus provided a more stringent test of potential group differences.

Quality Analysis

A quality analysis was conducted to determine the appropriateness of the selected studies for this meta-analysis. All of the studies included in this analysis reflected nonoverlapping empirical research projects published in

Table 1. Description of included regular past tense studies.

Study	Language	N			Age (years;months)			Scoring	Method
		SLI	TDA	TDY	SLI	TDA	TDY		
Eadie et al. (2002)	English	8		8	5;3		3;3	% correct –ed	C
Krantz & Leonard (2007)	English	15		15	5;2		3;4	% correct –ed	E
Leonard et al. (1992)	English	10	10	10	4;8	4;6	3;1	% correct –ed	EC
Leonard et al. (2003)	English	12	12	12	5;6	5;5	3;6	% correct –ed	E
Leonard et al. (2007)	English	16	10	15	5;1	5;0	3;5	% correct –ed	E
Marchman et al. (1999)	English	31	31		8;5	8;2		# omitted –ed	E
Marshall & van der Lely (2006)	English	14		27	12;03		7;6	% correct –ed ^a	E
Oetting & Horohov (1997)	English	11	11	11	6;3	6;3	3;11	# correct –ed	E
Proctor-Williams & Fey (2007)	English	13		13	7;10		5;6	% correct –ed	C
Redmond (2003)	English	7	7	7	6;1	6;0	3;7	% correct –ed	E
Redmond (2005)	English	10	13		6;7	6;7		% correct –ed	E
Rice & Wexler (1996)	English	37	45	40	4;10	5;0	3;0	% correct –ed	E
Rice et al. (2004)	English	130	117		6;0	5;11		% correct –ed	E
van der Lely & Ullman (2001)	English	11		12	11;2		5;9	% omitted –ed ^b	E
Windsor et al. (2000)	English	19	19	18	11;5	11;6	8;11	% omitted –ed	N
Hansson & Leonard (2003)	Swedish	14	14	10	4–5	4–5	2–3	% correct RP	E
Hansson & Nettelbladt (1995)	Swedish	5		5	5;5		4;6	% omitted RP	C
Hansson et al. (2000)	Swedish	14	14	14	4–5	4–5	2–3	% correct RP	C
Hansson et al. (2007)	Swedish	12		20	6;8		5	% correct RP	E
Lum & Bleses (2012)	Danish	13	20		7;8	7;11		% correct RP ^c	E
Rispens & De Bree (2014)	Dutch	36	39	40	8;1	7;9	5;4	% correct RP	E
Verhoeven et al. (2011)	Dutch	11	9		7;7	7;3		% omitted RP	N

Note. SLI = specific language impairment; TDA= typically developing age-matched; TDY = typically developing younger; C = conversational sample; E = elicitation probes; EC = conversation with elicitation probes; N = story retelling task; RP = regular past tense marking.

^aExperiment 2 monomorphemically legal cluster verbs for grammatical SLI (G-SLI) and combined language ability (LA) groups. ^bUnmarked high frequency regular verbs for G-SLI and LA1. ^cRegular small class verbs.

peer-reviewed journals. As discussed above, the researchers used professionally accepted participant identification and data collection methods. Furthermore, the majority of the researchers reported high reliability in data coding and analysis. Based on these factors, the studies included constitute an appropriate basis for judging the state of the evidence.

Because in studies of this type children cannot be randomly assigned to groups, we also examined group equivalence. With the exception of the slight age difference

favoring children with SLI over TDA children already noted above, the comparison groups were very similar on the variables on which they were chosen to be similar. In terms of other variables, children with SLI, as a group, had lower nonverbal IQ scores than their TDA counterparts. Across eight studies that used comparable nonverbal IQ tests, children with SLI scored a mean of 101.38 ($SD = 4.72$), whereas the TDA children scored a mean of 107.75 ($SD = 3.69$). Such a difference appears to reflect the SLI

Table 2. Description of included irregular past tense studies.

Study	Language	N			Age (years;months)			Scoring	Method
		SLI	TDA	TDY	SLI	TDA	TDY		
Leonard et al. (2003)	English	12		12	5;6		3;6	% marked past ^a	E
Marchman et al. (1999)	English	31	31		8;5	8;2		# unmarked past	E
Marshall & van der Lely (2012)	English	13		36	13;5		7;1	% marked past ^b	E
Redmond (2005)	English	10	13		6;7	6;7		% marked past	E
Rice et al. (2000)	English	21	23	20	4;8	4;11	3;0	% marked past	E
Rice et al. (2004)	English	130	117		6;0	5;11		% marked past	E
van der Lely & Ullman (2001)	English	11		12	11;2		5;9	% unmarked past ^c	E
Hansson & Nettelbladt (1995)	Swedish	5		5	5;5		4;6	% unmarked past	C
Vang Christensen & Hansson (2012)	Danish	11	11	11	6;4	6;5	4;2	% marked past	E

Note. SLI = specific language impairment; TDA= typically developing age-matched; TDY = typically developing younger; % marked past = irregular verbs marked for past tense; E = elicitation probes; # unmarked past = number of irregular past tense as bare stems; % unmarked past = past tense marking omitted in obligatory context; C = conversational sample.

^aIrregular scores available for SLI and TDY groups only. ^bTense marking score for grammatical SLI (G-SLI) and combined language ability (LA) groups. ^cUnmarked high frequency irregular verbs for G-SLI and LA1.

and TD populations more generally (Leonard, 2014, p. 20). This state of affairs leads researchers to choose between closely matching SLI and TDA groups in nonverbal IQ and running the risk that the SLI group is unrepresentative of the more general population of children with SLI, or including children who are representative of their respective groups with the result of slightly higher nonverbal IQ scores on the part of the TDA group (Leonard, 2014, p. 165). Most researchers have taken the second approach. The rationale for this selection method appears reasonable and such studies were included in the meta-analysis.

Reliability

Studies were coded for number of children per participant group, ages of the children, group identification methods, matching techniques, type of past tense elicitation task, scoring method, and results. A second trained coder working in the lab independently coded all of the identified studies for participant age, sample size, mean, and standard deviation of the target measures, resulting in 289 points of comparison with the original coder. On the first pass, the two coders agreed on 76% (219/289) of the points of comparison. Most of the discrepancies were due to the second coder reporting data for incorrect participant groups (e.g., a group with hearing impairment that was also included in the study instead of the TDA group, $n = 21$ comparison points), incorrect past tense measure (e.g., a combined past tense measure instead of separate measures for regular and irregular past tense or irregular past tense with over-regularizations excluded, $n = 31$ comparison points), and incorrect reporting of participant ages (e.g., incorrect conversion to years;months, $n = 7$ comparison points). A second, independent pass following additional coder training resulted in 96% agreement (278/289 points of comparison). The remaining 11 discrepancies were resolved by consensus. Six of the discrepancies occurred through errors in hand calculations of the means and standard deviations of the reported data. The remaining discrepancies were due to data input errors by one of the coders.

Effect Size Computation

Group means and standard deviations were extracted from eligible studies and entered into Comprehensive Meta-Analysis Software (CMA version 2; Borenstein, Hedges, Higgins, & Rothstein, 2005). For some studies, means and standard deviations had to be calculated from the raw data provided in the study. For one study, the F score and p value were converted and entered into CMA as Cohen's d . All data were then converted to Hedges's g effect sizes (Hedges, 1981). Hedges's g takes sample size into account and corrects for bias in small samples.

Heterogeneity

To uncover potentially unidentified moderator variables, we applied the measure I^2 , a measure of inconsistency across studies. I^2 describes the proportion of between-studies

variation resulting from heterogeneity instead of chance (Higgins, Thompson, Deeks, & Altman, 2003). In cases of high levels of heterogeneity, additional moderator variables were evaluated. Due to the low number of studies for irregular past, I^2 was restricted to the regular past tense data.

Results

Language type (English, non-English Germanic) was used as a moderator variable when calculating combined and overall effect sizes. A random-effects model was used to calculate combined effect sizes within each language type for regular past tense analyses. A fixed-effect model was then used to calculate an overall combined languages effect size. The variance across language type was assumed to differ and was not pooled in the overall calculations (Borenstein et al., 2005). The random-effects model assumes heterogeneity across studies within the analysis. Due to this heterogeneity, the true effect may vary due to between- as well as within-study variance (Borenstein, Hedges, Higgins, & Rothstein, 2009). Considering the variability in age at time of testing and data collection techniques, it is reasonable to assume that the true effect varies across studies in this meta-analysis. A random-effects model is, therefore, most appropriate. Because computation of the combined overall effect sizes involved only two effect sizes (that of English and of non-English Germanic), a fixed-effect model was applied in this case.

Due to the small number of included irregular past tense analyses, a random-effects model would likely yield an imprecise estimate of between-studies variance. In such cases, Borenstein et al. (2009) suggest using a fixed-effect model as a possible solution. Combined effect sizes for irregular past tense analyses were, therefore, calculated using a fixed-effect model.

Figures 3–6 provide effect sizes, measured in Hedges's g , for studies examining past tense production in children with SLI and their TD peers. The rectangular boxes on the corresponding forest plot represent the Hedges's g for each study. Negative effect sizes indicate poorer performance by children with SLI than by TD children. The solid horizontal lines indicate 95% confidence intervals (CIs). Lines crossing the zero line indicate studies with no significant differences between the SLI and TD groups. The diamonds represent combined effect sizes for English-only, non-English Germanic only, and overall performance across all studies. The center of the diamond represents the combined effect sizes, and the left and right edges represent their 95% CIs.

Regular Past Tense

When language type was considered as a moderator variable, there was a large overall effect size for regular past tense production by children with SLI versus their TDA peers, $g = -1.801$ ($SE = .235$, 95% CI $[-2.262, -1.339]$), which was significantly different than zero, $Z = -7.651$, $p < .0001$ (see Figure 3 and Table 3). The I^2 value was 90% (95% CI $[85\%, 93\%]$) suggesting high between-study heterogeneity.

Figure 3. Regular past tense production in specific language impairment versus typically developing age-matched children. The rectangular boxes represent Hedges's g for each study. The solid horizontal lines indicate 95% confidence intervals. The diamonds represent combined effect sizes for English-only, non-English Germanic only, and overall performance across all studies. The center of the diamond represents the combined effect sizes and the left and right edges represent their 95% confidence intervals.

The combined effect size was large for English studies, $g = -1.898$ ($SE = .392$, 95% CI $[-2.667, -1.129]$) and non-English Germanic studies, $g = -1.746$ ($SE = .294$, 95% CI $[-2.322, -1.169]$). The effect sizes for both language groups were significantly different than zero, $Z = -4.838$, $p < .0001$ and $Z = -5.935$, $p < .0001$, for English and non-English Germanic, respectively. The I^2 value was high for English studies ($I^2 = 92\%$, 95% CI $[87\%, 95\%]$) and moderate to high for Germanic studies ($I^2 = 62\%$, 95% CI $[0\%, 86\%]$). In all cases, children with SLI had significantly more difficulty than their TDA peers on regular past tense production.

There was a large overall effect size for regular past tense production by children with SLI versus their younger TD peers, $g = -1.180$ ($SE = .113$, 95% CI $[-1.401, -0.959]$), which was significantly different from zero, $Z = -10.47$, $p < .0001$ (see Figure 4 and Table 4). The I^2 value was 41% (95% CI $[0\%, 68\%]$) suggesting moderate between-study heterogeneity. The combined effect size was large for English, $g = -1.014$ ($SE = .151$, 95% CI $[-1.310, -0.717]$) as well as non-English Germanic studies, $g = -1.388$ ($SE = .169$, 95% CI $[-1.720, -1.057]$). The effect sizes for both language groups were significantly different from zero, $Z = -6.697$,

Figure 4. Regular past tense production in specific language impairment versus typically developing younger children. The rectangular boxes represent Hedges's g for each study. The solid horizontal lines indicate 95% confidence intervals. The diamonds represent combined effect sizes for English-only, non-English Germanic only, and overall performance across all studies. The center of the diamond represents the combined effect sizes and the left and right edges represent their 95% confidence intervals.

$p < .0001$ and $Z = -8.211$, $p < .0001$, for English and non-English Germanic, respectively. The I^2 value was moderate for English studies ($I^2 = 43\%$, 95% CI $[0\%, 70\%]$) and low to moderate for Germanic studies ($I^2 = 0\%$, 95% CI $[0\%, 74\%]$). Children with SLI in both language groups had significantly more difficulty producing regular past tense verbs than their TDY peers.

Due to the moderate to high levels of heterogeneity in the regular past tense analyses, elicitation method (conversation, elicitation, and narrative) and age group (preschool: 4;0–5;11, young school-age: 6;0–7;11, and older school-age: 8;0 and older) were also analyzed as moderator variables. Both the TDA and TDY comparisons had large effect sizes for conversation (TDA: $g = -1.618$, 95% CI $[-2.281, -0.954]$; TDY: $g = -0.953$, 95% CI $[-1.351, -0.555]$) and elicitation (TDA: $g = -1.998$, 95% CI $[-2.713, -1.284]$; TDY: $g = -1.227$, 95% CI $[-1.573, -0.882]$) and medium effect sizes for narratives (TDA: $g = -0.797$, 95% CI $[-1.320, -0.274]$; TDY: $g = -0.69$, 95% CI $[-1.34, -0.040]$). There were large effect sizes for all age groups in both the TDA (Pre/K: $g = -2.171$,

Figure 5. Irregular past tense production by specific language impairment versus typically developing age-matched children. The rectangular boxes represent Hedges's g for each study. The solid horizontal lines indicate 95% confidence intervals. The diamonds represent combined effect sizes for English-only, non-English Germanic only, and overall performance across all studies. The center of the diamond represents the combined effect sizes and the left and right edges represent their 95% confidence intervals.

95% CI [-3.261, -1.082]; young school-age: $g = -1.698$, 95% CI [-2.18, -1.216]; older school-age: $g = -1.176$, 95% CI [-1.484, -0.867]) and TDY (Pre/K: $g = -1.007$, 95% CI [-1.399, -0.614]; young school-age: $g = -1.227$, 95% CI [-1.659, -0.794]; older school-age: $g = -1.272$, 95% CI [-1.595, -0.948]) comparisons.

Although both English and non-English Germanic languages revealed large combined effect sizes, an inspection of the individual studies revealed that the two language types were not identical in the types of errors that occurred.

Figure 6. Irregular past tense production in specific language impairment versus typically developing younger children. The rectangular boxes represent Hedges's g for each study. The solid horizontal lines indicate 95% confidence intervals. The diamonds represent combined effect sizes for English-only, non-English Germanic only, and overall performance across all studies. The center of the diamond represents the combined effect sizes and the left and right edges represent their 95% confidence intervals.

In the non-English Germanic languages, productions of a present tense inflection occurred quite frequently, in some instances as frequently as infinitives. For example, Rispens and De Bree (2014) found that Dutch-speaking children with SLI produced forms such as *bakt* ("bakes") in place of the correct *bakte* ("baked") more frequently than infinitive forms such as *bakken* ("to bake"). These two types of errors constituted approximately 38% and 18% of their productions in regular past tense contexts, respectively. Bare stems (e.g., *bak*) were also produced by the children with SLI in this study (representing approximately 10% of their production in these contexts). The latter forms are appropriate in present tense first person singular contexts but were produced in past tense third person singular contexts here. The TDY children produced a much smaller number of present tense forms and no bare stems in regular past tense contexts. In their study on Swedish, Hansson and Leonard (2003) also found that present tense forms replaced past tense forms (e.g., *leker* "plays" in place of *lekte* "played") more frequently than did the infinitive (e.g., *leka* "to play" in place of *lekte*). Present tense forms and infinitive forms represented 54% and 42% of the errors, respectively. In English, the great majority of errors were productions of bare stems. However, Marchman, Wulfeck, and Ellis Weismer (1999) and Redmond (2003) employed an additional error category (described as "nonvalid" productions and "commission errors," respectively) that reportedly included productions of present tense third person singular forms in past tense contexts. The frequency of these errors was not specified.

Irregular Past Tense

There was a large overall effect size for tense marking in irregular past tense verbs by children with SLI versus their TDA peers, $g = -0.805$ ($SE = .106$, 95% CI [-1.013, -0.597]), which was significantly different from zero, $Z = -7.599$, $p < .0001$ (see Figure 5 and Table 5). The combined effect size was large for English studies, $g = -0.758$ ($SE = .109$, 95% CI [-0.971, -0.546]), and significantly different from zero, $Z = -6.987$, $p = .003$. The combined effect size for non-English Germanic studies also had a large effect size, $g = -1.746$ ($SE = .487$, 95% CI [-2.702, -0.791]), and differed significantly from zero, $Z = -3.583$, $p < .0001$. In all cases, children with SLI had significantly more difficulty than their TDA peers marking tense on irregular verbs.

There was a large overall effect size for irregular past tense production by children with SLI versus their TDY peers, $g = -0.828$ ($SE = .162$, 95% CI [-1.145, -0.511]), which was significantly different from zero, $Z = -5.121$, $p < .0001$ (see Figure 6 and Table 6). When language type was considered as a moderator variable, the combined effect size was large for non-English Germanic studies, $g = -1.333$ ($SE = .373$, 95% CI [-2.064, -0.603]) and medium for English studies, $g = -0.711$ ($SE = .180$, 95% CI [-1.063, -0.359]), with overlapping CIs between the two language groups. The effect sizes for both language groups were

Table 3. Specific language impairment (SLI) versus typically developing age-matched (TDA) regular past tense outcomes.

	Study name	Language	Hedges's <i>g</i>	95% confidence interval	
				Upper limit	Lower limit
1	Leonard et al. (1992)	English	-3.212	-4.514	-1.910
2	Leonard et al. (2003)	English	-2.694	-3.779	-1.609
3	Leonard et al. (2007)	English	-3.004	-4.123	-1.885
4	Marchman et al. (1999)	English	-0.519	-1.019	-0.019
5	Oetting & Horohov (1997)	English	-2.822	-3.981	-1.664
6	Redmond (2003)	English	-1.827	-3.019	-0.635
7	Redmond (2005)	English	-1.282	-2.159	-0.406
8	Rice & Wexler (1996)	English	-3.219	-3.874	-2.565
9	Rice et al. (2004)	English	-0.463	-0.715	-0.210
10	Windsor et al. (2000)	English	-0.659	-1.299	-0.019
	English combined		-1.898	-2.667	-1.129
11	Hansson & Leonard (2003)	Swedish	-1.864	-2.733	-0.994
12	Hansson et al. (2000)	Swedish	-1.059	-1.830	-0.288
13	Lum & Bleses (2012)	Danish	-2.810	-3.771	-1.849
14	Rispens & De Bree (2014)	Dutch	-1.977	-2.526	-1.428
15	Verhoeven et al. (2011)	Dutch	-1.076	-1.984	-0.169
	Non-English Germanic combined		-1.746	-2.322	-1.169
	Overall combined		-1.801	-2.262	-1.339

significantly different from zero, $Z = -3.578$, $p < .001$ and $Z = -3.960$, $p < .001$, for non-English Germanic and English, respectively. In this small group of studies, then, children with SLI had at least somewhat more difficulty marking irregular verbs for past tense than their TDY peers.

Although irregular past was scored as correct if the children produced either the appropriate irregular form or an overregularization, the latter had importance given that some of the alternative accounts of past tense difficulty in SLI are incompatible with productions of this type, as such productions seem to require application of a rule. For both English and non-English Germanic languages, many instances of overregularization were seen. For example,

Redmond (2005) noted that English-speaking children occasionally produced errors such as *fallen* in place of *fell*, as did Leonard et al. (2003), Marchman et al. (1999), Rice, Wexler, Marquis, and Hershberger (2000), and Rice, Tomblin, Hoffman, Richman, and Marquis (2004). Such errors were also seen in Danish (Vang Christensen & Hansson, 2012), as in *synkede* “sunked” in place of *sank* (“sank”). Although we included the Swedish study of Hansson, Nettelbladt, and Leonard (2000) and the Dutch study of Rispens and De Bree (2014) only in the analysis of regular past tense, in both of these studies, the authors also reported overregularization of irregular verbs (as in *springde* “runned” in place of *sprang* “ran” in Swedish).

Table 4. Specific language impairment (SLI) versus typically developing younger (TDY) regular past tense outcomes.

	Study name	Language	Hedges's <i>g</i>	95% confidence interval	
				Upper limit	Lower limit
1	Eadie et al. (2002)	English	-0.596	-1.545	0.354
2	Krantz & Leonard (2007)	English	-0.234	-0.932	0.465
3	Leonard et al. (1992)	English	-1.244	-2.168	-0.320
4	Leonard et al. (2003)	English	-1.668	-2.573	-0.763
5	Leonard et al. (2007)	English	-2.025	-2.876	-1.173
6	Marshall & van der Lely (2006)	English	-1.325	-2.020	-0.630
7	Oetting & Horohov (1997)	English	-0.993	-1.849	-0.137
8	Proctor-Williams & Fey (2007)	English	-1.015	-1.809	-0.221
9	Redmond (2003)	English	-1.223	-2.304	-0.143
10	Rice & Wexler (1996)	English	-0.546	-0.997	-0.095
11	van der Lely & Ullman (2001)	English	-1.306	-2.180	-0.432
12	Windsor et al. (2000)	English	-0.690	-1.340	-0.040
	English combined		-1.014	-1.310	-0.717
13	Hansson & Leonard (2003)	Swedish	-1.326	-2.195	-0.458
14	Hansson & Nettelbladt (1995)	Swedish	-0.837	-2.016	-0.341
15	Hansson et al. (2000)	Swedish	-0.976	-1.740	-0.213
16	Hansson et al. (2007)	Swedish	-1.655	-2.462	-0.848
17	Rispens & De Bree (2014)	Dutch	-1.593	-2.105	-1.080
	Non- English Germanic combined		-1.388	-1.720	-1.057
	Overall combined		-1.180	-1.401	-0.959

Table 5. Specific language impairment (SLI) versus typically developing age-matched (TDA) irregular past tense outcomes.

	Study name	Language	Hedges's <i>g</i>	95% confidence interval	
				Upper limit	Lower limit
1	Marchman et al. (1999)	English	–1.239	–1.777	–0.701
2	Redmond (2003)	English	–1.577	–2.493	–0.661
3	Rice et al. (2000)	English	–2.460	–3.235	–1.684
4	Rice et al. (2004)	English	–0.412	–0.664	–0.161
	English combined		–0.758	–0.971	–0.546
5	Vang Christensen & Hansson (2012)	Danish	–1.746	–2.702	–0.791
	Non-English Germanic combined		–1.746	–2.702	–0.791
	Overall combined		–0.805	–1.013	–0.597

Publication Bias

Publication bias may occur when only studies with significant effects in the expected direction are selected for publication. This type of bias within a meta-analysis data set may lead to overinflated combined effect sizes. A funnel plot, which plots effect size against standard error, allows for a visual identification of gaps in the data set due to potential publication bias (see Borenstein et al., 2009, and Cooper, Hedges, & Valentine, 2009, for a detailed explanation). The Duval and Tweedie (2000) trim and fill method provides an unbiased effect size estimate by trimming away studies causing asymmetry in the funnel plot, imputing hypothetical studies to create symmetry, and using the imputed data to recalculate the overall effect size. For the regular past tense studies in this meta-analysis, after trimming five studies, the adjusted effect size for combined comparison groups ($g = -1.091$, 95% CI $[-1.398, -0.785]$) was slightly smaller than the observed effect size ($g = -1.335$, 95% CI $[-1.65, -1.02]$), but remained in the medium to high range. Applying the trim and fill method to the irregular past tense studies, no trimming was required; thus, there was no change in the observed effect size (for both observed and adjusted, $g = -0.749$, 95% CI $[-0.932, -0.566]$).

Sensitivity Analysis

Sensitivity analyses of both regular and irregular past tense data were conducted to analyze the affect of each study on the overall effect size significance (Borenstein

et al., 2009). In this analysis, one study was removed and overall effect sizes were recalculated. The study was then reinserted, another study was removed, and effect sizes were once again recomputed. This process continued until each study had been removed. The resulting combined effect sizes ranged from -1.948 to -1.053 for regular past tense and -1.136 to -0.926 for irregular past tense studies—values that did not differ appreciably from those just reported. Thus, no single study had a major affect on the significance of the overall effect size.

Discussion

The purpose of this study was to evaluate differences in past tense production of children with SLI and their TD peers across Germanic languages. Regular past tense production was much less accurate in children with SLI than their age-matched and younger TD peers. Tense marking on irregular past tense verbs was also less accurate in children with SLI than both their age-matched and younger TD peers.

Several limitations in this study should be considered when interpreting the findings. Each of these limitations had the potential of affecting the combined effect sizes observed. It is possible that alternative search terms or databases could have led to different results. Certainly the inclusion of only those studies written in English limited the number of studies considered in each language, and, possibly, the number of different non-English Germanic languages

Table 6. Specific language impairment (SLI) versus typically developing younger (TDY) irregular past tense outcomes.

	Study name	Language	Hedges's <i>g</i>	95% confidence interval	
				Upper limit	Lower limit
1	Leonard et al. (2003)	English	–0.362	–1.141	0.417
2	Marshall & van der Lely (2012)	English	–0.680	–1.318	–0.042
3	Rice et al. (2000)	English	–0.790	–1.415	–0.166
4	van der Lely & Ullman (2001)	English	–1.029	–1.872	–0.187
	English combined		–0.711	–1.063	–0.359
5	Hansson & Nettelbladt (1995)	Swedish	–1.312	–2.570	–0.053
6	Vang Christensen & Hansson (2012)	Danish	–1.344	–2.241	–0.447
	Non-English Germanic combined		–1.333	–2.064	–0.603
	Overall combined		–0.828	–1.145	–0.511

included. Results might have also differed if more of the studies that we examined had met our inclusionary criteria. However, if the results changed after including such studies, they would not have changed dramatically, because the studies that were excluded for technical reasons produced results that were in keeping with the findings reported here. This was true not only for excluded studies in languages already included in the analysis, such as Dutch (e.g., de Jong, 1999) but also for excluded studies in languages not represented in the included data set, such as Norwegian (e.g., Simonsen & Bjerkan, 1998). For example, although de Jong (1999) calculated percentages of use of past tense forms by combining contexts for regular and irregular verbs, the differences between the SLI ($M = 77\%$) and both younger ($M = 98\%$) and same-age ($M = 99\%$) TD groups were notable. Furthermore, as in the Rispen and De Bree (2014) study on Dutch, de Jong (1999) reported that the children with SLI used present tense forms and bare stems as well as infinitives in contexts requiring past tense. In the Simonsen and Bjerkan (1998) study on Norwegian, the children with SLI were treated as individual case studies. However, when compared with larger groups of TD children at various ages, the children with SLI used both regular and irregular past tense forms at levels resembling TD groups who were approximately 2 years younger. Thus, although the findings of our meta-analyses provide only an estimate of the true state of affairs, there is no reason to believe that they are unrepresentative of the larger data set.

The method of data collection used in the included studies may also have influenced the results. Two of the three studies of regular past tense with CIs crossing zero (both comparisons of SLI and TDY groups; see Figure 4) employed conversational speech samples rather than elicitation tasks. On the other hand, two other studies involving the same languages (English, Swedish) that employed conversational samples did, in fact, reveal differences in line with the remaining studies. It seems possible that conversational speech samples are less discriminating because children have greater control of the vocabulary and the sentence constructions to use, enabling them to make greater use of well-practiced (and possibly memorized) linguistic forms. These observations notwithstanding, our moderator variable analyses revealed large effect sizes for studies involving conversation as well as elicitation methods, for SLI–TDY as well as SLI–TDA comparisons.

A wide age range was reflected in the studies included in the meta-analysis. Two of the studies with the oldest age groups (Marchman et al., 1999; Windsor, Scott, & Street, 2000) showed the smallest effect sizes. This might be expected given the fact that children with SLI show gradual increases in past tense ability with age, and may eventually reach ceiling levels. The remaining study with older children, that of van der Lely and Ullman (2001), showed a large effect size. However, for this study, the researchers deliberately selected those children with SLI who were already known to be weak in grammar. The large effect size for this study, then, may not have been representative of the differences ordinarily seen at older ages. Again, however,

moderator variable analyses revealed large effect sizes for all age groups.

Most of the studies included in the meta-analyses employed children with SLI who were clinically referred; the children then met additional criteria for inclusion. This raises the possibility that the findings are not applicable to all children who might meet criteria for SLI, but only those that drew the attention of family members or teachers. However, the study of Rice et al. (2004) employed children who had participated in the epidemiological study of Tomblin et al. (1997), and the children in the SLI group probably constituted a more representative sample of children with SLI in the general population. In this study, too, the children with SLI were less accurate than their peers, with CIs that never crossed zero.

In spite of the caution that should be exercised in interpreting particular details of this study, the results reported here add significant weight to the preliminary views that had been formed on the basis of the accumulating literature. Not only do children with SLI show less accuracy with past tense expression than their age mates, they are also less accurate than TD children who are younger in age. This finding holds for irregular as well as regular past tense, and for non-English Germanic languages as well as for English. Given the differences between children with SLI and younger children with typical development, it seems proper to conclude that past tense constitutes a special difficulty for children with SLI acquiring Germanic languages.

These findings have clear implications for evaluating the alternative accounts of past tense difficulties in children with SLI. First, the fact that irregular past showed large overall effect sizes when children with SLI were compared with both TDA and TDY children strongly suggests that the past tense weaknesses of children with SLI go beyond difficulty with brief inflections and the ability to segment inflections from stems. These findings indicate that both the accounts of Leonard et al. (1997) and Joannisse and Seidenberg (1998) are incomplete in their explanation of tense difficulties in SLI. The weaknesses with irregular past also suggest that past tense difficulties cannot be based entirely on operations that fall within procedural learning because irregular forms are assumed to be learned through the declarative memory system. Thus, the proposal of Ullman and Pierpoint (2005) seems incomplete as well.

The finding that regular past showed large overall effect sizes in the non-English Germanic languages is also suggestive that difficulties cannot be due entirely to the brevity of the morpheme. In these languages, past tense inflections are syllabic, which should present fewer challenges in both perception and production than the English consonantal inflections.

On the other hand, the phonetic characteristics of past tense inflections cannot be regarded as irrelevant. As a case in point, in the Leonard et al. (2003) study, English-speaking children with SLI scored much lower in their use of regular past tense than in their use of the phonetically identical passive participle inflections (e.g., *-ed* in *got kissed*). However, they were more likely to omit the latter than were

TDY children. Other studies (included in the meta-analyses) suggest that such details as the phonetic features of the stem or the phonotactic probability of the stem + inflection may also have an effect on these children's success with regular past tense inflections (Leonard et al., 2007; Marchman et al., 1999).

The frequent finding that children with SLI overregularized irregular verbs serves as evidence that much of their past tense use was the result of a creative, productive process. Such a finding seems to rule out an account that assumes a deficit in acquiring implicit rules, such as the one proposed by Gopnik and Crago (1991).

A problem based in dealing with computational complexity also seems insufficient. Along with the fact that weaknesses in the use of irregular past are not expected, this type of account should expect more difficulties with tense in non-English Germanic languages than in English. Danish, Dutch, and Swedish are "verb-second" languages. This means that when the sentence begins with an element other than the subject (e.g., a sentence beginning with the word "yesterday"), the verb occupies the second position of the sentence, before the subject. In the type of linguistic framework adopted in the computational complexity account (Chomsky, 1995; see van der Lely, 2004), a past tense verb in verb-second languages is assumed to either move from the syntactic position of verb to the position of INFL and then to the position of complementizer (COMP), or to have checking occur at each of these categories. English has no such requirement; past tense verb forms move to (or are checked at) INFL only. The number of linguistic operations assumed for tense, then, is different across languages, yet the cross-linguistic data are quite similar.

Based on the findings discussed thus far, the Extended Optional Infinitive account of Rice et al. (1998) and the input-misinterpretation account of Leonard et al. (2015) seem most consistent with the evidence. Both can account for overregularizations, an extended period of inconsistency with irregular as well as regular past tense forms, and errors that take the form of nonfinite verbs. However, one important detail in the findings runs counter to the predictions of both of these accounts. In these accounts, children's failure to express tense in a Germanic language is assumed to result in the use of a nonfinite form. Yet some studies, especially in the non-English Germanic languages, report present tense inflections as a frequent substitute. In some instances, task factors may have contributed to the children's use of present tense inflections. For example, in the Marchman et al. (1999) study, the children heard the present tense form immediately before being prompted to use the past tense form, as in: "This boy is walking. He walks every day. Yesterday, he ____?" It is plausible that the children simply repeated the form of the verb just heard instead of paying close attention to the past tense prompt. However, in other studies, this explanation would be less satisfactory. To elicit past tense use, Hansson and Leonard (2003) re-presented pictures that had been used for present tense earlier in the session, but, this time, presented them with the instructions (in Swedish) "You have seen this

picture before. Do you remember what the people did?" They then presented two pictures at a time, describing the first with past tense and prompting the child to respond to the second, as in *Mamma lagade mat och Kalle ____ (läste)* "Mommy cooked and Kalle ____ (read)." Yet present tense forms were frequently produced by the children.

To sum up, the alternative accounts of grammatical deficits in SLI are correct in focusing on past tense use, because children with SLI in these languages have, as this study shows, significant (and replicable) weaknesses in using past tense not only in English but in several other Germanic languages. However, each account has shortcomings, either by failing to predict difficulties with irregular past or regular past inflections that are syllabic in nature, by failing to credit these children with the ability to learn past tense rules, or by failing to account for errors in the form of inappropriate finite inflections rather than infinitives.

Although more work is needed to arrive at a fully satisfactory account of past tense difficulties in children with SLI, the strong evidence of serious weaknesses in this area shown by the meta-analyses provides important support to previous proposals that past tense problems might be a reliable clinical marker for SLI in English. It is important to note that the meta-analyses also indicate that the same could be true for other Germanic languages, in spite of the differences between these languages and English in the phonological details of past tense marking (syllabic versus consonantal) and in their rules of positioning finite verbs (verb second versus subject-verb-object). The findings of this study add further evidence that rather targeted measures of morphosyntactic proficiency may serve as a more appropriate means of identifying a language impairment than many currently used standardized tests (see Betz, Eickhoff, & Sullivan, 2013).

Unlike clinical markers in other fields (e.g., a physical characteristic that might suggest an underlying medical problem), the type of clinical marker identified here serves as more than a useful flagging tool. The very characteristic serving as a clinical marker—problems with past tense—should also be the focus of treatment. The weaknesses in past tense use are not only seen across studies in different Germanic languages; they are even apparent when younger TD children serve as the comparison group. They are, then, areas of extraordinary difficulty for children with SLI who are acquiring these languages. Clinicians should probably devote special attention to these problems in their service delivery, and clinical researchers should determine whether current methods of treating past tense weaknesses are sufficient or require significant modification.

Acknowledgments

Preparation of this review article was supported in part by Grants T32 DC00030 and R01 DC00458 from the National Institute on Deafness and Other Communication Disorders, awarded to the second author. We would also like to thank Jan de Jong for his insight and advice on non-English Germanic languages, as well as Julia Bergmann for her assistance with compliance reliability.

References

References marked with an asterisk indicate studies included in the meta-analysis.

- Bedore, L., & Leonard, L. (1998). Specific language impairment and grammatical morphology. *Journal of Speech, Language, and Hearing Research*, 41, 1185–1192.
- Bedore, L., & Leonard, L. (2001). Grammatical morphology deficits in Spanish-speaking children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, 44, 905–924.
- Bedore, L., & Leonard, L. (2005). Verb inflections and noun phrase morphology in the spontaneous speech of Spanish-speaking children with specific language impairment. *Applied Psycholinguistics*, 26, 195–225.
- Betz, S., Eickhoff, J., & Sullivan, S. (2013). Factors influencing the selection of standardized tests for the diagnosis of specific language impairment. *Language, Speech, and Hearing Services in Schools*, 44, 133–146.
- Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2005). *Comprehensive Meta-analysis, Version 2*. Englewood NJ: Biostat.
- Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2009). *Introduction to meta-analysis: Statistics in practice*. Chichester, United Kingdom: Wiley.
- Chomsky, N. (1995). *The minimalist program*. Cambridge, MA: MIT Press.
- Clahsen, H., Bartke, S., & Göllner, S. (1997). Formal features in impaired grammars: A comparison of English and German SLI children. *Journal of Neurolinguistics*, 10(2/3), 151–171.
- Conti-Ramsden, G., Botting, N., & Faragher, B. (2001). Psycholinguistic markers for specific language impairment (SLI). *Journal of Child Psychology and Psychiatry*, 42(6), 741–748.
- Cooper, H., Hedges, L. V., & Valentine, J. C. (Eds.). (2009). *The handbook of research synthesis and meta-analysis*. New York, NY: Russell Sage Foundation.
- de Jong, J. (1999). Specific language impairment in Dutch: Inflectional morphology and argument structure. *Groningen Dissertations in Linguistics*, 28.
- Duval, S., & Tweedie, R. (2000). A nonparametric “trim and fill” method of accounting for publication bias in meta-analysis. *Journal of the American Statistical Association*, 95(449), 89–98.
- *Eadie, F., Fey, M., Douglas, J., & Parsons, C. (2002). Profiles of grammatical morphology and sentence imitation in children with specific language impairment and Down Syndrome. *Journal of Speech, Language, and Hearing Research*, 45, 720–732.
- Gladfelter, A., & Leonard, L. (2013). Alternative tense and agreement morpheme measures for assessing grammatical deficits during the preschool period. *Journal of Speech, Language, and Hearing Research*, 56, 542–552.
- Gopnik, M., & Crago, M. (1991). Familial aggregation of a developmental language disorder. *Cognition*, 39, 1–50.
- Hansson, K. (1997). Patterns of verb usage in Swedish children with SLI: An application of recent theories. *First Language*, 17, 195–217.
- *Hansson, K., & Leonard, L. (2003). The use and productivity of verb morphology in specific language impairment: An examination of Swedish. *Linguistics*, 41(2), 351–379.
- *Hansson, K., & Nettelbladt, U. (1995). Grammatical characteristics of Swedish children with SLI. *Journal of Speech and Hearing Research*, 38, 589–598.
- *Hansson, K., Nettelbladt, U., & Leonard, L. (2000). Specific language impairment in Swedish: The status of verb morphology and word order. *Journal of Speech, Language, and Hearing Research*, 43(4), 848–864.
- *Hansson, K., Sahlén, B., & Mäki-Torkko, E. (2007). Can a “single hit” cause limitations in language development? A comparative study of Swedish children with hearing impairment and children with specific language impairment. *International Journal of Language & Communication Disorders*, 42(3), 307–323.
- Hedges, L. (1981). Distribution theory for Glass’ estimator of effect size and related estimators. *Journal of Educational Statistics*, 6(2), 107–128.
- Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *BMJ: British Medical Journal*, 327(7414), 557–560.
- Jacobsen, P., & Schwartz, R. (2002). Morphology in incipient bilingual Spanish-speaking preschool children with specific language impairment. *Applied Psycholinguistics*, 23, 23–41.
- Joanisse, M., & Seidenberg, M. (1998). Specific language impairment: A deficit in grammar or processing? *Trends in Cognitive Science*, 2, 240–247.
- *Krantz, L. R., & Leonard, L. B. (2007). The effect of temporal adverbials on past tense production by children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, 50(1), 137–148.
- Leonard, L. B. (2014). *Children with specific language impairment* (2nd ed.). Cambridge, MA: MIT Press.
- Leonard, L. B., & Bortolini, U. (1998). Grammatical morphology and the role of weak syllables in the speech of Italian-speaking children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, 41(6), 1363–1374.
- *Leonard, L., Bortolini, U., Caselli, M. C., McGregor, K., & Sabbadini, L. (1992). Morphological deficits in children with specific language impairment: The status of features in the underlying grammar. *Language Acquisition*, 2, 151–179.
- Leonard, L., & Deevy, P. (2011). Input distribution influences degree of auxiliary use by children with specific language impairment. *Cognitive Linguistics*, 22, 247–273.
- *Leonard, L., Deevy, P., Kurtz, R., Krantz Chorev, L., Owen, A., Polite, A., . . . Finneran, D. (2007). Lexical aspect and the use of verb morphology by children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, 50, 759–777.
- *Leonard, L., Deevy, P., Miller, C., Rauf, L., Charest, M., & Kurtz, R. (2003). Surface forms and grammatical functions: Past tense and passive participle use by children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, 46, 43–55.
- Leonard, L., Eyer, J., Bedore, L., & Grela, B. (1997). Three accounts of the grammatical morpheme difficulties of English-speaking children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, 40, 741–753.
- Leonard, L., Fey, M., Deevy, P., & Bredin-Oja, S. (2015). Input sources of third person singular –s inconsistency in children with and without specific language impairment. *Journal of Child Language*, 42, 786–820.
- Leonard, L., Hansson, K., Nettelbladt, U., & Deevy, P. (2004). Specific language impairment in children: A comparison of English and Swedish. *Language Acquisition*, 12(3/4), 219–246.
- Leonard, L. B., Sabbadini, L., Leonard, J. S., & Volterra, V. (1987). Specific language impairment in children: A cross-linguistic study. *Brain and Language*, 32(2), 233–252.
- *Lum, J., & Bleses, D. (2012). Declarative and procedural memory in Danish-speaking children with specific language impairment. *Journal of Communication Disorders*, 45, 46–58.
- *Marchman, V., Wulfeck, B., & Ellis Weismer, S. (1999). Morphological productivity in children with normal language

- and SLI: A study of the English past tense. *Journal of Speech, Language, and Hearing Research*, 42, 206–219.
- *Marshall, C. R., & van der Lely, H. K. (2006). A challenge to current models of past tense inflection: The impact of phonotactics. *Cognition*, 100(2), 302–320.
- *Marshall, C. R., & van der Lely, H. K. (2012). Irregular past tense forms in English: How data from children with specific language impairment contribute to models of morphology. *Morphology*, 22(1), 121–141.
- *Oetting, J., & Horohov, J. (1997). Past-tense marking by children with and without specific language impairment. *Journal of Speech, Language, and Hearing Research*, 40, 62–74.
- Paradis, J., & Crago, M. (2001). The morphosyntax of specific language impairment in French: An extended optional default account. *Language Acquisition*, 9(4), 269–300.
- *Proctor-Williams, K., & Fey, M. (2007). Recast density and acquisition of novel irregular past tense verbs. *Journal of Speech, Language, and Hearing Research*, 50, 1029–1047.
- *Redmond, S. (2003). Children's productions of the affix –ed in past tense and past participle contexts. *Journal of Speech, Language, and Hearing Research*, 46, 1095–1101.
- *Redmond, S. (2005). Differentiating SLI from ADHD using children's sentence recall and production of past tense morphology. *Clinical Linguistics & Phonetics*, 19(2), 109–127.
- Rice, M., Noll, K. R., & Grimm, H. (1997). An extended optional infinitive stage in German-speaking children with specific language impairment. *Language Acquisition*, 6(4), 255–295.
- *Rice, M., Tomblin, J. B., Hoffman, L., Richman, W. A., & Marquis, J. (2004). Grammatical tense deficits in children with SLI and nonspecific language impairment: Relationships with nonverbal IQ over time. *Journal of Speech, Language, and Hearing Research*, 47, 816–834.
- *Rice, M., & Wexler, K. (1996). Toward tense as a clinical marker of specific language impairment in English-speaking children. *Journal of Speech and Hearing Research*, 39, 1239–1257.
- Rice, M., Wexler, K., & Cleave, P. (1995). Specific language impairment as a period of extended optional infinitive. *Journal of Speech and Hearing Research*, 38, 850–863.
- Rice, M., 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, 1412–1431.
- *Rice, M., Wexler, K., Marquis, J., & Hershberger, S. (2000). Acquisition of irregular past tense by children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, 43, 1126–1145.
- *Rispen, J. E., & De Bree, E. H. (2014). Past tense productivity in Dutch children with and without SLI: The role of morphophonology and frequency. *Journal of Child Language*, 41(1), 200–225.
- Rothweiler, M., & Clahsen, H. (1994). Dissociations in SLI children's inflectional systems: A study of participle inflection and subject-verb agreement. *Scandinavian Journal of Logopaedics and Phoniatrics*, 18, 169–179.
- Simonsen, H. G., & Bjerkan, K. M. (1998). Testing past tense inflection in Norwegian: A diagnostic tool for identifying SLI children? *International Journal of Applied Linguistics*, 8(2), 251–270.
- Spelman, M., & Bol, G. (2012). The use of subject-verb agreement and verb argument structure in monolingual and bilingual children with specific language impairment. *Clinical Linguistics & Phonetics*, 26(4), 357–379.
- Thordardottir, E. (2008). Language-specific effects on task demands on the manifestation of specific language impairment of English and Icelandic. *Journal of Speech, Language, and Hearing Research*, 51, 922–937.
- 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(6), 1245–1260.
- Ullman, M., & Pierpoint, E. (2005). Specific language impairment is not specific to language: The procedural deficit hypothesis. *Cortex*, 41, 399–433.
- van der Lely, H. (2004). Evidence for and implications of a domain-specific grammatical deficit. In L. Jenkins (Ed.), *Variation and universals in biolinguistics* (pp. 117–144). Amsterdam: Elsevier.
- van der Lely, H. (2005). Domain-specific cognitive systems: Insight from Grammatical SLI. *Trends in Cognitive Sciences*, 9, 53–59.
- *van der Lely, H. K., & Ullman, M. T. (2001). Past tense morphology in specifically language impaired and normally developing children. *Language and Cognitive Processes*, 16(2–3), 177–217.
- *Vang Christensen, R., & Hansson, K. (2012). The use and productivity of past tense morphology in specific language impairment: An examination of Danish. *Journal of Speech, Language, and Hearing Research*, 55, 1671–1689.
- *Verhoeven, L., Steenge, J., & van Balkom, H. (2011). Verb morphology as clinical marker of specific language impairment: Evidence from first and second language learners. *Research in Developmental Disabilities*, 32, 1186–1193.
- Wexler, K. (2003). Lenneberg's dream: Learning, normal language development, and specific language impairment. In Y. Levy & J. Schaeffer (Eds.), *Language competence across populations: Toward a definition of specific language impairment* (pp. 11–61). Mahwah, NJ: Erlbaum.
- Wexler, K., Schaeffer, J., & Bol, G. (2004). Verbal syntax and morphology in typically developing Dutch children and children with SLI: How developmental data can play an important role in morphological theory. *Syntax*, 7(2), 148–198.
- Wexler, K., Schütze, C., & Rice, M. (1998). Subject case in children with SLI and unaffected controls: Evidence for the Agr/Tns Omission Model. *Language Acquisition*, 7, 317–344.
- *Windsor, J., Scott, C. M., & Street, C. K. (2000). Verb and noun morphology in the spoken and written language of children with language learning disabilities. *Journal of Speech, Language, and Hearing Research*, 43(6), 1322–1336.