Published in final edited form as:

Int J Speech Lang Pathol. 2015 December ; 17(6): 605–616. doi:10.3109/17549507.2015.1034174.

Quantifying the relative contributions of lexical and phonological factors to regular past tense accuracy

Amanda J. Owen Van Horne^{a,b} and Melanie Green Fager^c

^aDept. Of Communication Sciences & Disorders, University of Iowa

^bMember, DeLTA Center, University of Iowa

^cHarlem School District 122, Machesney Park, IL

Abstract

Purpose—Children with specific language impairment (SLI) frequently have difficulty producing the past tense. This study aimed to quantify the relative influence of telicity (i.e., the completedness of an event), verb frequency, and stem final phonemes on the production of past tense by school-age children with SLI and their typically-developing (TD) peers.

Method—Archival elicited production data from children with SLI between the ages of 6 and 9 and TD peers ages 4 to 8 were reanalyzed. Past tense accuracy was predicted using measures of telicity, verb frequency measures, and properties of the final consonant of the verb stem.

Result—All children were highly accurate when verbs were telic, the inflected form was frequently heard in the past tense, and the word ended in a sonorant/ non-alveolar consonant. All children were less accurate when verbs were atelic, rarely heard in the past tense, or ended in a word final obstruent or alveolar consonant. SLI status depressed overall accuracy rates, but did not influence how facilitative a given factor was.

Conclusion—Some factors that have been believed to be useful only when children are first discovering past tense, such as telicity, appear to be influential in later years as well.

Keywords

Past tense; telicity; lexical frequency; target selection; consonant cluster; child language; SLI

Clinicians working with children with language impairments are repeatedly faced with the question of how to effectively select stimuli for use in teaching a linguistic target. Appropriate target selection may enhance a client's success in therapy and improve intervention efficacy. However, the primary guidance for how to select targets to treat morphology comes from expert opinion (Crystal, 1985; Weiler, 2013) rather than data. Weiler, following Crystal, provides a series of recommendations for selecting verbs to teach past tense, including selecting frequent verbs that are acquired early, verbs with semantics

Contact Information: Amanda J. Owen Van Horne, 121 A SHC, Dept. of Communication Sciences & Disorders, 400 Hawkins Dr., University of Iowa, Iowa City, IA 52242. Phone: 319-335-6951. amanda-owen-vanhorne@uiowa.edu.

Declaration of interest: The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

that align with the meaning of past tense, and verbs with relatively simple word final phonological characteristics.

While there is considerable evidence supporting these suggestions, even a clinician well versed in this literature may struggle to balance these factors in order to select the best verbs for use in therapy. Assuming a clinician was adopting a developmental progression through these factors, what actual words should she select? Should she select a verb like *walk* because it is frequently heard in the past tense, even though the inflected verb ends in a potentially difficult to produce consonant cluster /kt/ and the event may not have a clear end point? Or a word like *answer*, which is somewhat abstract and less frequent, yet has an easy to produce word ending? Some attention has already been paid to how these factors influence learning in typical populations (Li & Shirai, 2001). However, the role that these factors should play in therapy is unknown, in part because these factors have only been studied alone and no one has asked if one factor is more influential than another in terms of production accuracy in children with language impairments.

At the end of his tutorial, Weiler (2013) recommends developing reference materials, similar to the references available in phonology (e.g., *40,000 Selected Words: Organized by Letter, Sound, and Syllable*; Blockcolsky, Frazer, & Frazer, 1987) that ease the task of selecting target words. *40,000 Selected Words* doesn't recommend that clinicians select words in a particular way, but instead provides a neutral tool cataloguing words by relevant variables that enables clinicians to select the words that support therapies that come from a variety of perspectives. As a first step in a research program examining the role of exemplars in treatment efficacy, this paper aims to quantify the relative contribution of factors that have been shown to influence regular past tense accuracy individually, but have not been considered together: lexical frequency, telicity, and word final phonology¹. In what follows we explore each of these areas of potential influence on accuracy in more detail.

Lexical Frequency

Weiler (2013) suggests choosing familiar verbs that are acquired early, a recommendation that could translate to selecting verbs that are frequent in the input (Marchman, Wulfeck, & Ellis Weismer, 1999; Oetting & Horohov, 1997). Single-route accounts of regular past tense acquisition emphasize the role of frequency in particular, arguing that children gradually accumulate enough types and tokens of each instance of past tense and use that information to extract patterns (e.g., McClelland & Patterson, 2002). Within this approach, high frequency words have stronger representations and are accessed more often and thus are more rapidly retrieved. Competition between different verb forms and the relationship between type and token frequency influences the way these patterns are constructed and thereby production accuracy (Albright & Hayes, 2003; Bybee & Slobin, 1982).

¹This paper is only concerned with the acquisition of regular verbs. Although, as will be described later, data from irregular verbs were available, the frequency differences between regular and irregular verbs were so substantial that it seemed statistically irresponsible to combine the data in the same regression model. Even studies that set out to directly test these two accounts often use data sets in which the most common regular verbs are still less common than the least common irregulars (van der Lely & Ullman, 2001). For this reason, we restrict our focus to the regular verbs.

Frequency can be characterized in two ways – overall frequency of the lexical item and frequency of the inflected form. Most studies have relied on frequency of the inflected form (Marchman, et al., 1999; Oetting & Horohov, 1997), but overall frequency is more readily available through resources like CHILDES (MacWhinney, 2000). Alternatively, combining these two values into a single ratio may effectively capture the competition between the inflected form and the bare stem, its nearest competitor. While other factors (e.g., telicity) may influence competition, this is most relevant to frequency because this particular information could potentially be collapsed into a single variable.

Higher frequency words do tend to be inflected more accurately, a finding that has been established across a variety of morphemes and studies, including plural (Rice & Oetting, 1993) and past tense (Marchman, et al., 1999; Oetting & Horohov, 1997, but see van der Lely & Ullman who found no frequency effects for TD children ages 5-9). Work with bilingual and ESL populations suggests that frequency influences morphological production across the lifespan (Blom & Paradis, 2013; Nicoladis, Palmer, & Marentette, 2007; Wulff, Ellis, Roemer, Bardovi-Harlig, & LeBlanc, 2009).

In studies of children with SLI, these children had greater difficulty inflecting low frequency words than their TD peers (Marchman, et al., 1999; Oetting & Horohov, 1997). Marchman et al. (1999) attempted to ameliorate the confound between frequency and regularity by doing median splits on frequency within the regular and irregular words and still found that low frequency regular words were more difficult for children with SLI. It could be that children with SLI have more trouble with low frequency forms because compensatory techniques like memorization work less well for these forms (van der Lely & Ullman, 2001) or it could be that these children have more difficulty identifying the statistical regularities associated with the patterns and sub-patterns of English past tense.

Lexical Aspect

Lexical aspect refers to the nature of an event in time (Vendler, 1957), including whether it is habitual, completed, ongoing, iterative, etc. In this paper, we will primarily be concerned with the contrast between telic events, which are events that have a clear endpoint (e.g., - *he ate a sandwich; he died yesterday*), and atelic events, which are events that are ongoing in nature (e.g., - *he was eating dinner; he's been dying all week long*). Most verbs can be grammatically realized as either telic or atelic events, but some are more naturally realized as a completed event while others are more naturally ongoing.

Selection of telic verbs are also recommended by Weiler (2013) because dynamic verbs with clear endpoints promote use of past tense in typical children (Leonard, Deevy, Kurtz, Krantz Chorev, Owen, Polite, Elam, & Finneran, 2007b). Some perspectives on morpheme acquisition consider the role of semantics in conjunction with frequency (Bloom, Lifter, & Hafitz, 1980; Goldberg, 2006) and assume that children must link the phonological form(s) of the morpheme to the semantic information it communicates (e.g., Tomasello, 2005). Telicity is most tightly associated with past tense; other meanings may be associated with other morphemes (e.g., statives tend to be associated with third person singular). The role that semantics can play in the initial acquisition and later generalization of morphological

marking is supported by data both from children and from computational models (Li & Shirai, 2001).

If one assumes that morphological markers, like words, have semantic content (Tomasello, 2005) then close alignment between the completed event described by the verb and the information contained by the simple past morpheme may help facilitate production of the morpheme. Indeed, initially, TD 2-year-olds produce simple past markers with telic events and progressive markers with atelic events. It has been argued that children discover the tense system by first attending to aspectual cues and then mapping that information on to tense (Bloom, Lifter, & Hafitz, 1980; Johnson & Fey, 2006; Shirai & Andersen, 1995). Lexical aspect also influences the likelihood that young children will choose a particular form of inflection. TD 3- to 4-year-old children were more likely to produce the simple past when a prototypical telic event was used (e.g., *drop*, *close*) than with a non-prototypical telic event (e.g., *rake*, *dance*; Leonard, et al., 2007b). However children with SLI were less sensitive than their TD peers to this contrast, using simple past equally often for both event types (Leonard, et al., 2007b).

Verbs that are highly telic are more likely to be heard in the past tense, introducing a confound between frequency and telicity. One possibility is that improved accuracy is entirely grounded in frequency: children are better at producing telic verbs in the past tense because they have heard the telic verb “*dropped*” more often than the atelic verb “*raked*”. Another possibility is that aspect is only a facilitative cue, initially helping children discover the semantics of the tense system. From this perspective, the influence of aspect may fade as children come to understand the meaning of the tense marker to be separate from the meaning of the verb. We explore this question by looking at whether frequency and telicity make unique contributions to children's accuracy at using regular past tense –ed and by examining accuracy in children who are older than the preschool years.

Phonology

Finally Weiler recommends selecting verbs based on their phonological characteristics. Phonological factors that influence accurate past tense production have received especially close examination in English due to the debate over the regular/irregular distinction. Even when only regular past tense is considered, phonology seems influential. There is general agreement that more complex contexts are more difficult (Leonard, et al., 2007a; Marchman et al., 1999; Marshall & van der Lely, 2007; Oetting & Horohov, 1997). Oetting and Horohov (1997) and Marshall and van der Lely (2007) found that children with SLI and their TD peers were less accurate when the addition of the past tense formed a consonant cluster. Marchman et al. (1999) suggested that children with SLI were more likely than their peers to omit past tense forms when the stem ended in an alveolar consonant in particular. They attributed this to problems with extracting the past tense pattern ~ that is, when a word ended in -t or -d the child was likely to assume that the word already had been inflected and not attempt to add more information. Words ending in -t and -d also take the syllabic allomorph, which is less frequent, but perhaps easier to produce than a consonant cluster. Theorists that assume that frequency of the phonological pattern drives acquisition would

attribute difficulties with the syllabic allomorph to its rarity rather than to the overlap between the phonological patterns of the stem and inflected forms (Albright & Hayes, 2003)

Relationship between lexical and phonological factors

There is preliminary evidence that both telicity and phonology contribute to past tense accuracy in TD children, but only when both are aligned in a non-facilitative way. Johnson and Morris (2007) elicited past tense forms from TD 2-year-olds and a single 4-year-old child with SLI using verbs that ended in an obstruent (e.g., *walk*, *jump*) or a non-obstruent (e.g., *roll*, *chew*) using elicitation scenarios that either had a clear endpoint (telic – *chewed up*) or were ongoing (atelic – *chewed all day*). TD children's overall accuracy dropped from around 50% to 31% when both the phonological and aspectual cues were non-facilitative (i.e., obstruent and atelic). There was no significant difference in performance between 2 facilitative factors (i.e., telic, non-obstruent) or 1 facilitative and 1 non-facilitative factor. In contrast, the results from the child with SLI suggested a more graded effect. This child was 100% accurate for telic items with a non-obstruent coda, 33% accurate when either one of those factors was present, and 0% accurate when neither factor applied. As only one child with SLI was studied, it is not clear whether this pattern generalizes to a larger group of children with SLI.

Study Purpose

To summarize, variable production of regular past tense seems to be mediated by the frequency, aspectual, and phonological properties of the individual lexical item being inflected. We examine children who are older than those included in previous studies to determine whether these factors continue to influence regular past tense production during the school years over and above the contribution of frequency. We also examine whether frequency, aspect, or phonology make a greater contribution to accuracy. We ask the following questions:

1. What frequency information best captures the influence of frequency on past tense accuracy?

We predict that a ratio of *Past Frequency* to *Stem Frequency* will best capture the influence of frequency measures on regular past tense accuracy since it is a single value that captures one source of competition between inflected and uninflected forms.

2. Does telicity uniquely contribute to past tense accuracy over and above frequency information?

If learning proceeds entirely based on distributional information, telicity should be subsumed under frequency. However, if the semantic information influences morpheme learning (Tomasello, 2005), telicity may make a unique contribution over and above that made by frequency, suggesting that attention to both factors is warranted².

²This may still be distributional information, but it would be information related to how the world and events distribute rather than information housed entirely within the language itself.

3. Does difficulty with obstruent or alveolar consonants influence accuracy?

The evidence is mixed with regard to phonology. Two possible interpretations seem available:

- a. If obstruent consonants account for most of the variance then phonological complexity would be the reason for difficulty with regular –ed inflections (Marshall & van der Lely, 2007).
- b. If alveolar consonants accounted for the majority of the variance we would argue for difficulty extracting the regular past tense pattern (Marchman, et al., 1999).

4. How do the frequency, telicity, and phonological measures contribute to accuracy when both are analyzed together?

Based on Johnson and Morris (2007), we predict that both lexical and phonological measures will contribute separately to accuracy, but we make no predictions about their relative contributions to accuracy (i.e., will the contributions be additive or multiplicative or, as Johnson and Morris found, will one dominate the other?).

Methods

Data Collection

Participants—The children studied here were initially recruited for participation in a study of complex syntax production (Owen, 2010). All research reported below was completed in accordance with the ethical guidelines for human subjects research as described in the Belmont report and required by the National Institutes of Health.

Fourteen children with SLI (5;0-8;1; M = 6;7, SD = 11.1 mos.) and 24 TD children participated in Owen (2010), providing the data reanalyzed in this study. The TD children were divided into 2 groups: one age-matched group (Age) and a second younger group matched on expressive vocabulary and mean length of utterance (MLU). The 13 age-matched TD children (M = 6;7; SD = 10.6 mos.) were matched within 3 months of age to a child with SLI. The age-matched group served to demonstrate how TD same-age children would perform on the tasks – that is, are the tasks age appropriate and would mastery be expected? The other 11 TD children (M = 4;4, SD = 3 mos.) were 4-year-olds who were matched on two measures. They scored within 5 raw score points on the *Expressive Vocabulary Test* (EVT, Williams, 1997) and within .35 words on MLU in a 100-utterance language sample to a child with SLI. This group served as an utterance length and vocabulary size control. The original study examined the role of complex syntax on past tense accuracy and both vocabulary size and utterance length have been linked to the ability to produce these types of sentences, thus both were considered as controls when selecting language-matched participants instead of just MLU as is common. This group served as an utterance length and vocabulary size control so that differences in past tense production could not be attributed to different levels of linguistic development

The children in the SLI group met at least two of the following four diagnostic criteria: a) enrolled in intervention within the last 12 months; b) scored below the 10th percentile on the *Structured Photographic Expressive Language Test – 2nd edition* (SPELT-II; Werner & Kresheck, 1983); c) scored below a standard score of 7 on the nonword repetition subtest of the *NEPSY* (Kemp, Korkman, & Kirk, 1998); d) obtained a composite standard score at or below 85 on the *Test of Narrative Language* (TNL, Gillam & Pearson, 2004). TD children had no history of intervention and scored at or above the typical range on all of the speech and language measures described above. All children scored above 83 on the *Kaufman Basic Intelligence Test-II, Matrices Subtest*, (KBIT-II, Kaufman & Kaufman, 2004). A cutoff score of 83 was used as, with the standard error on the test, this kept all children's "true" score reliably above a score of 80. All children also passed a hearing screening. All children were monolingual speakers of Standard American English and were recruited from college, farming, and manufacturing towns in the Midwest. All children were capable of producing word final /s, z, t/ and /d/ in single word contexts. This was assessed using a picture naming task, with 40 items devoted to /t/ and /d/ in one and two syllable words (e.g., *coat, rabbit, cloud, mermaid*). For items that the child did not know the name of or provided another name for (e.g., bunny for rabbit) the examiner named the item and asked the child to imitate her. Half of the items tested were clusters (e.g., *tent, dentist, hand, windshield*) but only a composite score of 80% or better on all t/d items was required in order to be included in the study. Table 1 summarizes relevant information about the participants and more information can be found in Owen (2010).

Stimuli—The data for this study are verbs that were elicited as a part of a study on complex syntax (Owen, 2010) using puppet show elicitation tasks (adapted from Crain & Thornton, 1998 and Redmond & Rice, 2001). These data were chosen because the same verbs were repeated across elicitation conditions, yielding multiple productions of the same verb for each child. Each condition targeted a different sentence type, consisting of a main clause made up of a subject and a verb followed by another finite clause: 1) two coordinated clauses, 2) a main clause with a finite complement clause and 3) a main clause and a temporal adverbial clause. Thirty-six (36) sentences were constructed for each sentence type. The protocols for eliciting the stimuli and the design of the sentences are summarized in Owen (2010). Verbs were systematically repeated across conditions, taking into consideration the linguistic restrictions of each clause type. With the exception of the main clause verbs of the finite complement condition, all verbs were repeated across multiple conditions, reducing the effect of the elicitation condition on the results reported here. Children altered the target lexical items routinely and thus, although verbs were elicited multiple times, the number of examples of each verb produced varied widely across children (see Appendix A).

Procedures—For each sentence type, children watched a short puppet show enactment and then a neutral prompt was posed (*Tell me about the show.*). The experimenter provided the first word of the first clause (*Ernie...*) and encouraged the child to complete the rest of the sentence (*Ernie wore a hat and Elmo jumped up and down*). If a child did not respond or provided an incomplete response, the examiner asked the child to say the whole sentence. If the child still did not produce a full response, the examiner provided the first clause of the

sentence (*Ernie wore a hat and...*) and again asked the child to complete the sentence. All responses were audio-recorded and transcribed, using the on-line written record as a guide. Transcription reliability was computed for nine sets of transcripts (three children from each diagnostic group) for word level accuracy (mean: 90%; range: 75-100%) and for tense marking (mean: 93% range: 81-100%).

Data Reduction

Outcome Variables—In Owen, 2010, only the two-clause responses were scored for past tense accuracy since the emphasis was on sentence level variables and responses from both regular and irregular verbs were employed. In this study, all regular verbs were analyzed regardless of whether they were produced in one- or two-clause responses since the focus is on word level variables. 8208 tokens (38 children $\times$ 3 conditions $\times$ 36 items $\times$ 2 clauses) were potentially available for analysis. Of these, 823 were one-clause responses that were not analyzed in Owen (2010) but are included here. A total of 4553 responses were discarded due to being irregular verbs ($n = 2175$), being non-target verbs ($n = 920$), using non-simple past inflections (e.g., past progressive, simple present tense, $n = 836$) or being unintelligible ($n = 652$). Verbs that were not the target for that item but matched another target verb were retained ($n = 413$). Thus the analyses involved 3625 tokens of 60 regular verbs.

Predictor Variables

Lexical frequency: Three lexical frequency variables were calculated from the parental lexical frequency information in the CHILDES database (MacWhinney, 2000): *Lemma Frequency*, *Past Frequency* and *Stem Frequency*. This database was selected because it is freely available; however, it does not differentiate between the noun and verb forms of the target item (e.g., *take a walk*, *we walk around*) or between participial and regular past tense forms (*I walked*, *he has walked*), which influenced the frequency counts below. All frequency variables were log transformed to normalize their distribution. See Appendix B for the transformed frequency values.

- *Lemma Frequency* was the sum of frequency information for all forms of the word. For example, for *walk*, the frequencies of *walk*, *walked*, *walks*, and *walking* were summed.
- *Past Frequency* only included correct productions of the simple past form (e.g., *walked*). Participial forms were indistinguishable from simple past (e.g., *the dog got walked*, *he has walked home*, *he walked*).
- *Stem Frequency* consisted of the uninflected stem forms, whether correctly produced (e.g., *I walk*), a noun form (*the walk*) or bare due to omission (e.g., *he walk*).

Past Frequency is the most commonly reported variable (e.g., Marchman et al., 1999; Oetting & Horohov, 2001). However, familiarity with the word, rather than with the word in its inflected form may be most critical. If so, then *Lemma Frequency* should be the best predictor. *Stem Frequency* reflects word familiarity but it also is a count of the form that is the most common error. Thus it may serve more as a competitor rather than a facilitator. We

also considered two derived variables, the *Ratio of Past to Stem* and the *Ratio of Past to Lemma*, because these ratios had the potential to capture the competition between inflected and alternate forms in a single value.

Telicity: Most previous studies have used lists of activity/achievement verbs to construct categories of prototypical telic and atelic verbs. As a study of archival data, this was not possible. Instead, we used adult ratings of the verbs situated within the elicitation scenarios. Seven adults read the elicitation scenarios and rated each verb on a scale of 1-4 as ongoing (1) or completed (4). All 7 raters used the full rating scale, but means (range: 2.2-3.4) and standard deviations (range: .92-1.38) varied across raters. Each rater's scores were normalized. Then z-scores were averaged across raters to yield a single *Telicity* score for each verb, as shown in Appendix B.

Phonological factors³: We coded the place and manner characteristics of the final consonants. For the manner variable, words that ended in [b, d, g, p, t, k, θ, f, s, ʃ, ð, v, z, ʒ, dʒ, tʃ] were coded as *Obstruent*; words ending in other sounds were coded as non-obstruent. In terms of place, homo-organic production of the final consonant of the stem and the past tense inflection depresses accuracy (Marchman, et al., 1999). Thus we also coded verbs ending in [d, t, s, ʃ, z, ʒ, dʒ, tʃ, j, r, l, n] as *Alveolar*. Verbs ending in all other sounds were coded as non-alveolar. Sample verbs, along with coding for phonology, are given in Appendix B. Syllabic allomorphs were not coded separately as this would have led to a grouping of only 6 verbs (*count, float, paint, plant, point, and rest*).

Statistical Approach

Following Baayen, Davidson, and Bates (2008) and Jaeger (2008), we adopted a mixed model logistic regression as our statistical approach. Although this approach is less transparent than ANOVA or simple linear regression, it allows one to carry out by-item and by-subject analyses simultaneously; it weights the reliability of a factor based on how many data points are available; and the binomial distribution mirrors the underlying distribution of data coded as correct or omitted. Categorical variables were contrast coded. Correlations among continuous variables are reported in Table 2. Given the widely differing scales, variables were centered and scaled to make comparison across parameter estimates more straightforward. Scaling the variables makes it possible to compare the magnitude of the parameter estimates without being distracted by the different units associated with each variable. This makes it possible to directly compare variables, one goal of the study.

³Although we would hypothesize that the phonotactic probability of the inflected forms would be an excellent potential predictor of accuracy and would provide a segment-level correlate to the frequency based measures used for the whole word, this measure is not reported for two reasons. First, neither the Vitevitch and Luce (2004) nor the Storkel and Hoover (2010) databases incorporate inflected forms into their dictionaries. This is misleading in terms of the weightings for the probability of those phoneme combinations or phonemes in that position. Even selecting only the last syllable or coda of the inflected word yielded unusual frequency metrics since attested – indeed frequent – combinations were not found in the database. Second, adding an inflection makes the word longer, shifting the word to a more or less common length (see Storkel, 2004). Thus the length normalization procedure does not work properly for words that are inflected forms. Following a reviewer's suggestion, we removed phonotactic probability entirely rather than report nonsignificant results when the metrics are in question.

Results

Mixed model regression includes both random and fixed effects. The random intercepts included were child and verb; the random slope included was elicitation condition ($X^2(5) = 184.40$, $p < 0.0001$). Random effects mathematically adjust the intercepts for each child and verb. For instance, the random intercept for child assumes that each child has a different mean accuracy level. Owen (2010) found that children were equally accurate at the first and second clauses in the coordinated condition. Compared to this condition, children were generally less accurate in the first clause of the adverbial condition and the second clause of the finite complement condition. The random slopes for condition were included in the model and they assume that each condition has a unique effect on performance trajectory. Finding significant fixed effects takes into account these corrections.

Group (Age, SLI, MLU) was a fixed effect in all models based on known group differences. As expected, both the MLU-matched group ($p = .055$) and the SLI group ($p < .0001$) were more likely to omit regular past tense than the age-matched group. Although the results for the MLU-matched group were marginal, this effect was included since it has the potential to influence the outcomes of the overall model. Models were compared via chi-squared analyses with and without random effects to ensure that random effects were necessary. Only the final model is reported (See Table 3).

We initially tested which lexical frequency and telicity measures accounted for past tense production over and above the contributions of group. Following our predictions that a single ratio value would capture competition between forms, we initially tested the use of a ratio of *Past Frequency* to *Lemma Frequency* or *Past Frequency* to *Stem Frequency*, but these models were not significant. Next, *Past Frequency*, the variable most highly related to the outcomes (see Table 2), was entered into the model, significantly improving model fit over the reduced model ($X^2(3) = 12.88$, $p = 0.005$). As *Past Tense Frequency* increased, so did accuracy for all three groups of children. Then *Lemma Frequency* and *Stem Frequency* were each regressed against *Past Frequency* and the residual values were entered⁴. The residuals of *Lemma Frequency* resulted in slightly lower log likelihood values in the model, $X^2(3) = 8.48$, $p = 0.037$. Addition of *Stem Frequency* residuals did not further improve model fit ($p > .50$). Testing for interactions demonstrated that the MLU-matched group was influenced by *Lemma Frequency* ($p < .002$), showing a slight improvement when words were more frequent overall. Figure 1 shows the influence of the *Past Frequency* (Fig 1a) and *Lemma Frequency* (Fig 1b) on the accuracy for each group.

We next regressed *Telicity* against the two frequency predictors and assessed whether the residuals contributed significantly to the outcomes. *Telicity* residuals were a significant predictor and did not interact with group ($p > .50$). Counter to our hypothesis that telicity and frequency would be interchangeable, this suggests that telicity contributes uniquely to past tense accuracy.

⁴Regressing one variable (e.g., *Lemma Frequency*) against another variable (e.g., *Past Frequency*) tells you about the relationship between these two variables. The residuals are a measure of the error in the model. Entering the residuals essentially provides information about how surprised you should be by the *Lemma Frequency* given that you know *Past Frequency*. It is useful when variables are highly correlated with each other.

Finally, the phonological factors (*obstruent* and *alveolar*) were entered into the model. The phonological factors improved model fit over the lexical factors alone ($X^2(6) = 29.94$, $p < .0001$) and vice versa ($X^2(3) = 34.11$, $p < .0001$) indicating that both were necessary to explain past accuracy. These results, illustrated in Figure 2, indicate both difficulties with phonological complexity and with application of the past tense pattern. The MLU-matched group was less influenced by the presence of a word final obstruent than the other two groups ($p = .03$).

A major goal of this study was to compare the relative contribution telicity, frequency, and phonology to past tense accuracy. Recall that the use of standardized beta coefficients allows the direct comparison of these values. Examination of the regression table (Table 3) shows that the phonological factors and lexical frequency/telicity factors both account for similar sized changes in the child's rate of inflection, but in opposite directions. Age and SLI groups do not differ with regard to the influence of frequency, telicity, and phonology, as evidenced by the lack of significant interactions. For an MLU-matched child, the effect of phonology is less than that observed for an age-matched child, and the influence of lemma frequency is greater.

Discussion—Although prior work has independently considered the influence of frequency, telicity, and stem-final phonology on past tense accuracy, we are not aware of work that has considered how these factors may together contribute to past tense accuracy. Contrary to our predictions, we found that past frequency, lemma frequency and telicity all affect the accuracy of children. Words ending in obstruents and alveolars are less likely to be accurate, suggesting that both phonological complexity and the relationship of stems to the phonological pattern of inflected words influence past tense accuracy. A key contribution of this study is that past frequency and telicity are not interchangeable variables and that these variables continue to influence accuracy well beyond the initial emergence of past tense markers. A second contribution is documentation that both frequency/telicity related variables and phonological variables make approximately equal contributions to overall accuracy.

For the age-matched children and for the children with SLI, the lexical and phonological coefficients were comparable in magnitude. More frequently inflected words, words depicting completed events, and words ending in non-obstruents and non-alveolars enhanced accuracy. Less frequently inflected words, word depicting ongoing events, and words ending in obstruents and alveolars reduced accuracy. Take, for instance, the words *close* and *kiss* which both have high telicity ratings (see Appendix B). The beneficial effect of telicity cancels out the detrimental effect of ending in /s/ or /z/. *Close* is slightly more telic and substantially more common in the past tense form and thus more accurate than *kiss*. Contrast this with *play*, another frequently correct word. *Play* is highly accurate primarily because it is phonologically simple; it is rated as atelic and is rarely heard in the past tense form. Verbs that are never observed in the past tense in CHILDES, *hum*, *rake*, and *giggle*, also get low telicity ratings. In these cases any advantages provided by the phonological form cannot overcome the negative influence of telicity and frequency. Finally, verbs, such as *imagine*,

exercise, and *rest*, which get low ratings on both telicity and frequency measures *and* have uniformly poor phonological characteristics, are among the least likely verbs to be inflected.

Replication with other morphemes would be valuable. One would imagine that the primary finding – that semantics, frequency, and phonology all make unique contributions to accuracy – is not unique to past tense, though the details might change if another morpheme paradigm were selected or if another language were examined. For instance, for third person singular, one would focus on how –s is realized and its role in clusters, the frequency that words are heard in that form, and the role of stative/habitual actions rather than telicity.

How to best capture frequency effects—The frequency of the inflected form is a well-examined factor (Marchman et al., 1999; Oetting & Horohov, 1997). Our finding that *Past Frequency* was a significant predictor of past tense accuracy is generally consistent with these results. While our results are generally in agreement with those of Marchman et al. (1999) in finding that the frequency of the inflected form influenced all groups similarly, we differ from Oetting and Horohov (1997) who only found group differences for verbs that were rarely inflected and from van der Lely and Ullman (2001) who also found that the children with SLI, but not the typical children, were sensitive to frequency effects. The ages of the participants, methods of elicitation, selection of stimuli and analyses differ considerably across these studies. One possible explanation could be that we employed a frequency continuum, rather than selecting verbs because they were either very high or very low frequency. Studies which have found no influence of frequency on past tense inflection in typical participants (e.g., van der Lely and Ullman, 2001) may have suffered from a restriction of range in the regular forms due to attempts to match the regular and irregular exemplars on frequency. In combination with the other published findings, these results would suggest that clinicians should ensure minimal familiarity with target words but realize that frequency in the inflected form influences past tense accuracy more than overall frequency.

The fact that the ratio of *Past Frequency* to *Lemma Frequency* could not replace the two frequency variables is consistent with the findings of Nicoladis, Palmer, and Marentette (2007) who found that both type and token frequency of the inflected forms uniquely influenced production of past tense by bilingual children. A minimum number of exposures to the target form seemed to be required; greater rates of exposure led to more accurate production. Others examining second language learners have also observed that *Past Frequency* and *Lemma Frequency* separately contribute to accuracy levels (Blom & Paradis, 2013). Second language learners may have less exposure to English and smaller verb vocabularies. This could explain the findings from the MLU-matched group as compared to the age-matched group. As seen in Figure 1, for the MLU-matched children, positive residual values are associated with increased accuracy, suggesting that greater familiarity with the word form, in any form, is associated with greater accuracy.

That said, the SLI and MLU-matched groups were also matched on EVT scores, yet only the MLU-matched group showed this interaction. Although the EVT score was selected with the goal of matching the groups on expressive vocabulary scores, it may not have accomplished this goal. It is well known that the same score on a standardized test may be obtained in

different ways. More detailed or stronger representations of words may support better inflection and such information is not captured by an offline standardized test like the EVT. Children with SLI are necessarily older and have more real-world experience than younger language matched peers. Follow up work is necessary to better understand this finding.

For the age and SLI groups, lemma frequency at first appears to work in competition with the past frequency, as we had hypothesized, but this effect is not reliable (p-values of .12 to .19) and should not be over-interpreted. Although nonsignificant, lemma frequency may be influential and worthy of consideration in future studies like this one.

Telicity and Past Frequency make unique contributions to accuracy—Prior work has suggested that telicity is primarily useful for helping children ‘crack into’ the tense system (e.g., Leonard et al., 2007b). Our results indicate that telicity influences the use of regular past tense beyond the preschool years (Johnson & Fey, 2006; Johnson & Morris, 2007; Leonard, et al., 2007b), suggesting a role for semantics that goes beyond that of initial discovery. Telic events may not just be a way to bootstrap into the inflection system, only to be superseded by frequency later on. Instead telicity may serve as a persistent cue, enhancing the online production of past tense across the lifespan (Wulff et al., 2009). Furthermore, instead of only classifying events categorically as telic or atelic, the rating system we used treated events as falling on a continuum. Even though adults’ intuitions about the “completed-ness” of events is highly related to the frequency with which these words are produced in past tense ($r=.30$; $p=.02$), both factors separately contribute to the likelihood that past tense will be produced. Clinicians may be able to promote initial accuracy by selecting highly telic events like those used in Johnson and Fey (2006) or Leonard et al. (2007b) or by presenting verbs in ways that make them appear more telic (see Appendix C for suggestions). The unique contributions of *Telicity* and *Past Frequency* are critical for the understanding what promotes accurate inflection.

The role of these semantic factors may vary with dialect. Though simple past appears to express completed-ness across a wide variety of English dialects, present perfect may be more commonly used in British and Australian English (Engel & Ritz, 2000; Hundt & Smith, 2009). Careful work identifying the semantic contexts that differentiate past perfect and simple past would be useful, just as extant work on the differences between past progressive and simple past (Leonard et al., 2007b) have provided insights into how children learn to mark tense and aspect.

Phonology—Both word final obstruents and alveolars influenced the use of past tense. As expected, *Obstruent* verbs (verbs ending in nasals, fricatives, affricates, and stops) were less accurate than words that ended in sonorants (liquids, glides, and vowels), reflecting problems forming consonant clusters (Marshall & van der Lely, 2007). Verbs ending in *Alveolars* were also difficult to a similar degree, as has been previously reported by Marchman et al. (1999).

Examination of the six verbs that ended in –t or –d and therefore took the syllabic allomorph indicated that these verbs tended to be inflected less than average for all three groups, but many of these verbs were also low frequency verbs making it difficult to draw conclusions

about phonology from such a small sample. This study is limited by the fact that it uses archival data – the groups were initially selected and matched for other purposes and the lexical items are not carefully controlled. To a certain extent, the fact that the lexical items were not carefully controlled makes it possible to examine the verbs as a continuum, rather than categorically, as others have done, but it also means that sampling may have been uneven, particularly with regard to the phonological variables. Replication of our results in other elicitation settings or in spontaneous samples would strengthen the findings.

Some challenges do arise in extension of our results to other populations. All of our speakers spoke standard American English. The phonological characteristics of past tense are different when past tense is produced in other dialects. For instance, African American English (AAE) often reduces or omits consonant clusters at the ends of words. That said, our finding that words that end with a vowel are more likely to receive past tense marking than words that end with an obstruent is consistent with other findings (e.g., Oetting & MacDonald, 2001). Further work with a variety of dialects would also be informative, both for examining phonology and for consideration of the role of telicity and other semantic cues.

Group Differences & Combined Factors—Although studies, including ours, consistently find that the children with SLI do more poorly than their age-matched TD peers in overall accuracy (i.e., Leonard et al., 2007a, b; Marchman et al., 1999; Oetting & Horohov, 1997; van der Lely & Ullman, 2001), whether or not this is exacerbated by a non-facilitative context was less clear. Our models suggest that each factor is influential, but that interactions with diagnostic group are not likely. Similarly, the frequency, telicity and phonological characteristics did not interact with each other, though we did test for this. Thus while each one made production more difficult in an additive fashion, they did not modify each other. Unlike Johnson and Morris (2007), who only observed differences when all of the non-facilitative factors were present in an item, we saw more gradual changes. Johnson and Morris only saw incremental changes in the SLI group. In our study, similar changes were observed across all groups. The primary areas in which group interactions were observed were in words ending in obstruent consonants and lemma frequency, both of which influenced the MLU-matched group differently than the age-matched group. Differences between the age-matched group and the SLI group were limited to main effects.

Our choice of statistical method may be influencing our outcomes. Logistical regression emphasizes differences between two extreme points and changes in the error rate. When errors are rare (as they are for the age-matched group), a factor that doubles the *error rate* (e.g., from 2% to 4%) may not lead to a great change in the *overall accuracy level* (e.g., from 98% to 96%). When errors are more common, a doubling in rate leads to a greater change in overall accuracy levels. Such differences may make it appear that an interaction is present in more traditional statistical approaches (e.g., ANOVA), which assess changes in accuracy, rather than changes in rate. In logistic regression, the factors may be acting in similar ways on the likelihood of producing an error across groups of children; it is actually the base rate of errors that differs.

Conclusions and Clinical Implications: Our results suggest that verbs that are telic, commonly inflected in the past tense, and end in facilitative phonological contexts is likely to lead to be produced accurately, while verbs that are atelic, less commonly inflected in the past tense and end in obstruents or alveolars are likely to be produced less accurately. Frequency and telicity continue to influence accuracy into the school-age years, a finding that suggests that attention to both pieces of information is necessary, just as attention to both the number of syllables and the number of phonemes are commonly taken into account in treatment of speech sound disorders.

Verb selection for teaching a target morpheme should not be random or arbitrary (Crystal, 1985; Weiler, 2013). One might assume that variables that exert a greater influence on accuracy would warrant greater attention from a clinician. This study does not provide a direct test of how to select items for treatment, but begins to provide information about which factors influence accuracy in the moment, a good first step.

In carrying out this study, we also obtained a rank ordering of 60 regularly inflected verbs from highly accurate to much less accurate. Many therapy approaches build on the assumption that understanding the difficulty of the stimulus items is critical. Many developmentally based approaches to speech and language therapy assume that beginning with highly stimulable, easy, or well known targets reduce processing load and make acquisition easier (Rvachew & Nowak, 2001; Weiler, 2013). Other approaches assume that beginning with carefully selected complex targets may lead to widespread generalization throughout the system (e.g., Geirut, 2001). In either case, the rank ordering may be valuable to clinicians who wish to consider the role of difficulty in producing a correct inflection in therapy. Continued research into which of these factors induce more rapid progress or greater gains in terms of generalization beyond the treated items is a critical need in the area of grammar treatment.

Acknowledgments

This work was supported by an internal research grant from the University of Iowa. We would like to thank the children and families who participated in this project and the support of Augustana College (Allison Haskill) and the Scottish Rite Program (Elizabeth Merrifield) for assistance with subject recruitment. This paper benefited from discussion within the Language Discussion Group at the University of Iowa. Susan Wagner Cook provided advice on statistical analysis. Jean Gordon read and commented on the paper and its presentation of material. Portions of the paper were presented at BUCLD held in Boston, MA in 2009.

Appendix A

The table below shows the number of times each of the verbs used in these analyses was elicited and produced. The children sometimes chose to substitute one verb for another. As mentioned in the text, if the child substituted an irregular verb (*made*) for a regular verb (*baked*) or an unstudied verb (*cheered*) for a target verb (*clapped*), the response was discarded. If the child substituted a target verb (*jumped*) for another target verb (*bounced*, *hopped*) the response was retained and analyzed as it was produced. In the complement clause condition, children strongly preferred to use the verbs *know* and *think*, a finding also reported by Kidd, Lieven, & Tomasello (2006). With the exception of the verb, *guess*, lexical selection was generally even across all three groups of children.

Clause/Condition	Verb Stem	Number Elicited Per Child	Number Produced (Percent Accuracy)		
			Age	SLI	MLU
Regular Main Complement Clause Verbs	answer	3	16 (100%)	15 (40%)	11 (100%)
	believe	3	28 (85%)	13 (31%)	20 (80%)
	discover	3	32 (97%)	13 (85%)	20 (100%)
	guess	6	60 (82%)	147 (77%)	41 (95%)
	imagine	3	20 (80%)	12 (58%)	16 (82%)
	remember	3	29 (97%)	14 (29%)	18 (100%)
Regular Main Coordinate/Temporal Clause Verbs	bake	2	20 (90%)	17 (59%)	15 (80%)
	clap	2	17 (88%)	18 (56%)	18 (78%)
	clean	2	18 (88%)	16 (56%)	17 (76%)
	climb	2	28 (96%)	24 (58%)	21 (71%)
	color	2	25 (96%)	29 (72%)	27 (81%)
	count	2	21 (81%)	21 (33%)	19 (68%)
	exercise	2	15 (66%)	17 (53%)	19 (63%)
	fish	2	16 (94%)	14 (57%)	15 (87%)
	float	2	18 (78%)	17 (29%)	14 (57%)
	giggle	2	16 (88%)	15 (33%)	10 (70%)
	growl	2	20 (90%)	15 (40%)	14 (64%)
	hum	2	19 (95%)	18 (44%)	15 (87%)
	listen	2	21 (86%)	16 (44%)	15 (67%)
	paddle	2	18 (78%)	15 (47%)	10 (60%)
	paint	2	21 (85%)	19 (63%)	14 (71%)
	plant	2	20 (75%)	18 (17%)	17 (76%)
	point	2	19 (95%)	19 (37%)	16 (56%)
	rake	2	23 (91%)	20 (70%)	14 (71%)
	rest	2	15 (40%)	14 (7%)	13 (31%)
	roll	2	20 (90%)	19 (84%)	17 (82%)
	sail	2	18 (88%)	16 (50%)	15 (87%)
	scratch	2	17 (82%)	17 (47%)	15 (67%)
	slip	2	24 (100%)	26 (61%)	20 (90%)
	snore	2	17 (88%)	17 (65%)	7 (71%)
	stamp	2	20 (95%)	20 (60%)	16 (81%)
	stir	2	14 (92%)	10 (80%)	14 (71%)
	trip	2	23 (96%)	21 (71%)	19 (89%)
	turn	2	18 (89%)	14 (57%)	12 (83%)
	walk	2	22 (86%)	15 (73%)	18 (83%)
	wave	2	19 (89%)	19 (42%)	18 (83%)
	whisper	2	18 (94%)	20 (50%)	17 (88%)
	whistle	2	18 (100%)	19 (63%)	15 (87%)
	wiggle	2	19 (89%)	21 (43%)	14 (57%)
	work	2	15 (93%)	14 (64%)	11 (91%)

Clause/Condition	Verb Stem	Number Elicited Per Child	Number Produced (Percent Accuracy)		
			Age	SLI	MLU
Regular Second Clause Verbs in All Conditions	yawn	2	22 (95%)	21 (57%)	17 (76%)
	yell	2	21 (95%)	22 (68%)	16 (69%)
	bark	3	20 (90%)	18 (89%)	18 (78%)
	bounce	3	13 (100%)	5 (40%)	11 (82%)
	carry	3	22 (95%)	16 (75%)	12 (83%)
	close	3	18 (83%)	12 (75%)	15 (60%)
	cough	3	30 (90%)	23 (61%)	21 (86%)
	crawl	3	20 (100%)	14 (64%)	11 (100%)
	cry	3	23 (91%)	18 (89%)	20 (80%)
	dance	3	28 (100%)	19 (47%)	24 (92%)
	hop	3	18 (89%)	14 (57%)	15 (93%)
	hug	3	23 (87%)	25 (72%)	16 (94%)
	jump	3	50 (98%)	49 (88%)	42 (90%)
	kiss	3	35 (89%)	23 (74%)	20 (85%)
	play	3	18 (100%)	19 (63%)	24 (83%)
	scare	3	33 (100%)	25 (76%)	26 (88%)
	smile	3	27 (100%)	21 (67%)	18 (67%)
	sneeze	3	30 (97%)	24 (58%)	22 (86%)
	squish	3	25 (92%)	17 (71%)	14 (71%)
	stretch	3	36 (92%)	23 (65%)	26 (73%)

Appendix B

Phonological codes and unstandardized frequency and telicity values for each verb along with the predicted percent correct use from the regression model in Table 3 for the SLI group. Verbs are presented from easiest/most accurate (*close*) to hardest/least accurate (*exercise*).

Verb Stem	Manner	Place	Telicity	Log Past Frequency	Log Lemma Frequency	SLI predicted percent correct
close	Obstruent	Alveolar	3.92	2.39	2.99	0.87
play	Other	Other	1.71	1.28	2.33	0.86
scare	Other	Alveolar	3.00	2.22	2.42	0.86
answer	Other	Alveolar	3.75	1.15	2.21	0.85
jump	Obstruent	Other	2.75	2.01	2.87	0.84
carry	Other	Other	2.29	1.46	2.61	0.84
cry	Other	Other	1.79	1.94	2.99	0.83
slip	Obstruent	Other	3.31	1.45	1.99	0.83
trip	Obstruent	Other	3.44	1.04	2.21	0.83
walk	Obstruent	Other	2.19	2.10	3.06	0.82

Verb Stem	Manner	Place	Telicity	Log Past Frequency	Log Lemma Frequency	SLI predicted percent correct
remember	Other	Alveolar	2.33	1.68	3.30	0.82
kiss	Obstruent	Alveolar	3.42	1.60	2.68	0.82
climb	Obstruent	Other	2.38	1.95	2.55	0.81
discover	Other	Alveolar	3.25	0.90	1.30	0.80
hug	Obstruent	Other	2.92	1.00	2.09	0.79
yell	Other	Alveolar	2.63	1.34	2.02	0.79
crawl	Other	Alveolar	2.96	0.90	1.89	0.79
stamp	Obstruent	Other	2.81	1.26	1.52	0.79
sneeze	Obstruent	Alveolar	3.54	1.08	1.64	0.78
cough	Obstruent	Other	2.96	0.60	2.00	0.78
stretch	Obstruent	Alveolar	2.50	2.09	2.21	0.77
guess	Obstruent	Alveolar	2.88	1.08	2.99	0.76
work	Obstruent	Other	1.44	1.84	3.32	0.76
color	Other	Alveolar	1.50	1.68	3.13	0.76
point	Obstruent	Alveolar	2.81	2.34	3.74	0.76
turn	Obstruent	Alveolar	1.69	2.43	3.45	0.76
bake	Obstruent	Other	2.13	1.34	2.02	0.76
roll	Other	Alveolar	1.56	1.57	2.64	0.75
hop	Obstruent	Other	2.13	1.04	2.24	0.75
smile	Other	Alveolar	2.13	0.85	2.19	0.74
whistle	Other	Alveolar	2.14	0.85	1.91	0.73
clean	Obstruent	Alveolar	1.94	1.83	2.85	0.73
count	Obstruent	Alveolar	2.56	1.00	2.67	0.73
stir	Other	Alveolar	2.31	0.48	1.89	0.73
paint	Obstruent	Alveolar	2.06	0.00	1.46	0.72
wave	Obstruent	Other	2.00	0.85	1.93	0.72
plant	Obstruent	Alveolar	2.31	1.62	2.68	0.72
believe	Obstruent	Other	1.96	0.60	2.41	0.71
bark	Obstruent	Other	2.04	0.70	1.76	0.71
wiggle	Other	Alveolar	1.56	1.32	1.58	0.71
clap	Obstruent	Other	2.00	0.70	1.81	0.71
squish	Obstruent	Alveolar	2.42	1.11	1.48	0.69
bounce	Obstruent	Alveolar	2.25	1.04	1.93	0.69
yawn	Obstruent	Alveolar	2.87	0.30	1.66	0.69
snore	Other	Alveolar	1.81	0.70	1.23	0.68
whisper	Other	Alveolar	1.81	0.48	1.72	0.68
scratch	Obstruent	Alveolar	2.00	1.20	2.08	0.68
dance	Obstruent	Alveolar	1.79	1.20	2.64	0.68
float	Obstruent	Alveolar	2.38	0.70	1.65	0.67
growl	Other	Alveolar	1.88	0.00	1.40	0.65
listen	Obstruent	Alveolar	1.44	2.43	4.22	0.64

Verb Stem	Manner	Place	Telicity	Log Past Frequency	Log Lemma Frequency	SLI predicted percent correct
sail	Other	Alveolar	1.44	0.48	1.49	0.64
hum	Obstruent	Other	1.44	0.00	2.44	0.63
paddle	Other	Alveolar	1.69	1.04	2.94	0.63
rake	Obstruent	Other	1.75	0.00	1.41	0.63
giggle	Other	Alveolar	1.69	0.00	1.23	0.63
fish	Obstruent	Alveolar	1.56	0.30	2.97	0.61
imagine	Obstruent	Alveolar	1.75	0.00	1.90	0.57
rest	Obstruent	Alveolar	1.25	0.00	2.65	0.55
exercise	Obstruent	Alveolar	1.50	0.00	1.75	0.55

Appendix C: Tips for eliciting telic events

While public databases do not exist for telicity rankings, the following strategies may increase the perceived completed-ness of an event:

- have a clear endpoint to the event
 - finish coloring a picture in completely.
 - arrange chairs to be jumped over in a line, rather than a circle
 - crawl to a destination rather than aimlessly
- make the event punctual rather than iterative/durative
 - hug/kiss a single person rather than several
 - close a single door completely and rapidly
 - bark once rather than for a long time
- make the event countable
 - plant a specific number of plants in pots rather than planting things in rows in a garden – run out of pots at the end
- make the event short, less than 6 seconds
 - fish – and catch a fish – quickly
 - make the coloring picture small so it can be completed rapidly
 - climb a short ladder rather than a tall ladder

References

- Albright A, Hayes B. Rules vs. analogy in English past tenses: A computational/experimental study. *Cognition*. 2003; 90(2):119–161. [PubMed: 14599751]
- Baayen RH, Davidson DJ, Bates DM. Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*. 2008; 59:390–412.
- Blockcolksky V, Frazer JM, Frazer DH. 40,000 Selected Words Organized by Letter, Sound, and Syllable. San Antonio: Psychological Corporation; 1987.

- Blom E, Paradis J. Past tense production by English second language learners with and without language impairment. *Journal of Speech, Language, and Hearing Research*. 2013; 56:281–294.
- Bloom L, Lifter K, Hafitz J. Semantics of verbs and the development of verb inflection in child language. *Language*. 1980; 56:386–412.
- Bybee JL, Slobin DI. Rules and schemas in the development and use of the English past tense. *Language*. 1982;265–289.
- Crain S, Thornton R. *Investigations in Universal Grammar A Guide to Experiments on the Acquisition of Syntax and Semantics*. Cambridge, MA: The MIT Press; 1998.
- Crystal D. Some early problems with verbs. *Child Language Teaching and Therapy*. 1985; 1:46–53.
- Engel DM, Ritz MEA. The use of the present perfect in Australian English. *Australian Journal of Linguistics*. 2000; 20(2):119–140.
- Gierut JA. Complexity in Phonological Treatment: Clinical Factors. *Language, Speech, and Hearing Services in Schools*. 2001; 32(4):229–241.
- Gillam R, Pearson N. *Test of Narrative Language: Examiner's Manual*. Austin: PRO-ED; 2004.
- Goldberg A. *Constructions at Work*. Oxford: Oxford University Press; 2006.
- Hundt M, Smith N. The present perfect in British and American English: has there been any change, recently? *ICAME journal*. 2009; 33:45–63.
- Jaeger TF. Categorical data analysis: Away from ANOVAs transformation or not, and towards logit mixed models. *Journal of Memory and Language*. 2008; 59:434–446. [PubMed: 19884961]
- Johnson BW, Fey ME. Interaction of lexical and grammatical aspect in toddlers' language. *Journal of Child Language*. 2006; 33:419–435. [PubMed: 16826833]
- Johnson BW, Morris S. Clinical implications of the effects of lexical aspect and phonology on children's production of the regular past tense. *Child Language Teaching and Therapy*. 2007; 23:287–306.
- Kaufman AS, Kaufman NL. *KBIT-2: Kaufman Brief Intelligence Test*. San Antonio: Pearson; 2004.
- Kemp S, Korkman M, Kirk U. *Essentials of NEPSY Assessment*. New York: Harcourt Assessments; 1998.
- Kidd E, Lieven EV, Tomasello M. The acquisition of complement clause constructions: A sentence repetition study. *Cognitive Development*. 2006; 21:93–107.
- Leonard LB, Davis J, Deevy P. Phonotactic probability and past tense use by children with specific language impairment and their typically-developing peers. *Child Language Teaching and Therapy*. 2007a; 21:747–758.
- Leonard LB, Deevy P, Kurtz R, Krantz Chorev L, Owen A, Polite E, Elam D, Finneran D. Lexical aspect and the use of verb morphology by children with specific language impairment. *Journal of Speech, Language, and Hearing Research*. 2007b; 50:759–777.
- MacWhinney B. *The Childes Project: Tools for Analyzing Talk*. 3rd. Psychology Press; 2000. 8083rd Edition 2 Volume Set with CD-Rom
- Marchman VA, Wulfeck B, Ellis Weismer S. Morphological productivity in children with normal language and SLI: A study of the English past tense. *Journal of Speech, Language, and Hearing Research*. 1999; 42:206–219.
- Marshall CR, van der Lely HKJ. The impact of phonological complexity on past tense inflection in children with Grammatical-SLI. *International Journal of Speech-Language Pathology*. 2007; 9:191–203.
- McClelland JL, Patterson K. Rules or connections in past-tense inflections: What does the evidence rule out? *Trends in cognitive sciences*. 2002; 6(11):465–472. [PubMed: 12457897]
- Nicoladis E, Palmer A, Marentette P. The role of type and token frequency in using past tense morphemes correctly. *Developmental Science*. 2007; 10:237–254. [PubMed: 17286847]
- Oetting JB, Horohov JE. Past-tense marking by children with and without specific language impairment. *Journal of Speech, Language, and Hearing Research*. 1997; 40:62–74.
- Owen AJ. Factors affecting accuracy of past tense production in children with specific language impairment and their typically-developing peers: The influence of verb transitivity, clause location, and sentence type. *Journal of Speech, Language, and Hearing Research*. 2010; 53:993–1014.

- Pinker S, Ullman MT. The past and future of the past tense. *Trends in cognitive sciences*. 2002; 6(11): 456–463. [PubMed: 12457895]
- Pinker S, Prince A. Regular and irregular morphology and the psychological status of rules of grammar. In: Lima SD, Corrigan R, Iverson GK, editors *The Reality Of Linguistic Rules*. Amsterdam: John Benjamins; 1994. 321–352.
- Plunkett K, Juola P. A connectionist model of English past tense and plural morphology. *Cognitive Science*. 1999; 23(4):463–490.
- Redmond S, Rice ML. Detection of irregular verb violations by children with and without SLI. *Journal of Speech and Hearing Disorders*. 2001; 44:655–669.
- Rice ML, Oetting JB. Morphological deficits of children with SLI: evaluation of number marking and agreement. *Journal of speech and hearing research*. 1993; 36:1249–1257. [PubMed: 8114492]
- Rvachew S, Nowak M. The effect of target-selection strategy on phonological learning. *Journal of speech, language, and hearing research*. 2001; 44(3):610–623.
- Shirai Y, Andersen RW. The acquisition of tense-aspect morphology: A prototype account. *Language*. 1995; 71:743–762.
- Storkel HL. Methods for minimizing the confounding effects of word length in the analysis of phonotactic probability and neighborhood density. *Journal of Speech, Language, and Hearing Research*. 2004; 47:1454–1468.
- Storkel HL, Hoover JR. An online calculator to compute phonotactic probability and neighborhood density on the basis of child corpora of spoken American English *Behavior research methods*. 2010; 42:497–506. [PubMed: 20479181]
- van der Lely HKJ, Ullman UM. Past tense morphology in specifically language impaired and normally developing children. *Language and Cognitive Processes*. 2001; 16:177–217.
- Vendler Z. Verbs and times. *The Philosophical Review*. 1957; 66:143–160.
- Vitevitch M, Luce P. A web-based interface to calculate phonotactic probability for words and nonwords in English. *Behavior Research Methods, Instruments, & Computers*. 2004; 36:481–487.
- Weiler B. Verb selection and past-tense morphology: Crystal's criteria revisited. *Topics in Language Disorders*. 2013; 33:152–164.
- Werner E, Kresheck J. *SPELT-II: Structured Photographic Expressive Language Test*. 2nd. DeKalb, IL: Janelle Publications; 1983.
- Williams KT. *Expressive Vocabulary Test*. Circle Pines, MN: American Guidance Service; 1997.
- Wulff S, Ellis N, Roemer U, Bardovi-Harlig K, Leblanc C. The acquisition of tense–aspect: Converging evidence from corpora and telicity ratings. *The Modern Language Journal*. 2009; 93:354–369.

Figure 1.

The influence of past tense frequency (Fig 1a) and lemma frequency (Fig 1b) on past tense accuracy for the three groups of children, age-matched and MLU matched typically developing children and children with SLI. Past and lemma frequency values are reported in quartiles for the purposes of illustration.

Figure 2.

The influence of place and manner of the stem final consonant on past tense accuracy for the three groups of children, age-matched and MLU matched typically developing children and children with SLI..

Summary of participants' performance on standardized assessments given as means and (standard deviations)

Table 1

Assessment	Score Type	Normal Range	Group		
			AGE	MLU	SLI
SPELT-II	Percentiles	16-84	72.4 (17.8)	81.2 (16.7)	10.22 (8.9)
TNL Index	SS	85-115	109.5 (16.3)	-	83.3 (14.9)
NEPSY: NWR	SS	7-13	12.3 (2.6)	-	7.3 (1.8)
KBIT-2	SS	85-115	99.6 (9.37)	106.2 (8.1)	97.4 (11.97)

Note: SS = standard score. The Normal Range is the mean +/- 1 SD and is given because the scales shift across tests. Assessments included the Structured Photographic Expressive Language Test – II (SPELT-II), the Test of Narrative Language (TNL), the NEPSY Nonword repetition Subtest (NEPSY-NWR) and the Kaufman Basic Intelligence Test –II, matrices subtest (KBIT-II). Scores listed are the mean for each participant group: Age-matched (AGE), vocabulary and utterance length matched (MLU) and specific language impairment (SLI). Numbers in parentheses are the standard deviations. The MLU-matched group did not receive the TNL or the NEPSY because the norms did not include 4 year olds.

Table 2
Correlations between untransformed variables of interest

		<u>Past</u>	<u>Lemma</u>	<u>Stem</u>
Log Lemma Frequency	<i>r</i>	0.65*	--	--
	<i>p</i>	<.0001	--	--
Log Stem Frequency	<i>r</i>	0.51*	0.96*	--
	<i>p</i>	<.0001	<.0001	--
Telicity	<i>r</i>	0.30*	-0.08	-0.09
	<i>p</i>	0.02	0.53	0.47

Table 3

Full Model predicting the accuracy of past tense production using standardized variables.

Random Effects:		Variance	Std.Dev.		
	Verb	0.31	0.55		
	Subject (Intercept; ref = Adverb)	4.65	2.16		
	Coord	2.81	1.68		
	Subord	2.04	1.43		
Fixed effects:					
	Estimate	Std. Err	Z-score	p value	
	Intercept	2.82	0.26	10.83	<0.0001
	SLI	-1.16	0.27	-4.29	<0.0001
	MLU	-0.54	0.28	-1.94	0.052
	Log Past Freq	0.44	0.13	3.52	<0.001
	Res. Log Lemma Freq	-0.15	0.09	-1.55	0.12
	Res. Telicity	0.39	0.10	3.76	<0.001
	SLI * Res. Log Lemma Freq	0.08	0.06	1.31	0.19
	MLU * Res. Log Lemma Freq	0.19	0.06	2.97	0.003
	Obstruent	-0.42	0.10	-4.20	<0.0001
	Alveolar	-0.30	0.10	-3.17	0.002
	SLI * Obstruent	0.09	0.06	1.44	0.15
	MLU * Obstruent	0.13	0.07	2.06	0.04

N = 3625, Verbs = 60, Subjects = 38

AIC = 2699, BIC = 2817, Log Likelihood = 1331, Deviance = 2661

Note: The random effects include conditions from Owen, 2010. Coord = Coordinate Condition, Subord = Subordinate Condition, with the Adverbial condition serving as the reference condition (ref = Adverb). Children with specific language impairment (SLI) and their language matched counterparts (MLU) are compared with the Age-matched group, which serves as the reference point included in the intercept. Freq = frequency. Res = Residual. Akaike information criteria (AIC) and Bayesian information criteria (BIC) are reported and were used to evaluate the goodness of the model along with log likelihood comparisons.