Developmental Psychology
2010, Vol. 46, No. 6, 1497–1504

0012-1649/10/$12.00 DOI: 10.1037/a0020668

Children's Judgments of Regular and Irregular Novel Past-Tense Forms: New Data on the English Past-Tense Debate

Ben Ambridge
University of Liverpool

Is language governed by formal rules or by analogy to stored exemplars? The acquisition of the English past tense has long played a central role in this debate. In the present study, children rated the acceptability of a regular and an irregular past-tense form of each of 40 novel verbs (e.g., *fleeped*, *flept*) using a 5-point scale. The novel verbs were chosen to vary continuously along the orthogonal dimensions of (a) similarity to existing regular forms and (b) similarity to existing irregular forms. A developmental progression was observed whereby the acceptability of novel regulars was shown to increase as a function of similarity to existing regulars, with the magnitude of this effect increasing with age. The acceptability of novel irregulars was shown to increase as a function of similarity to existing irregulars, with no developmental changes observed. These findings are discussed in the light of 3 current models of past-tense acquisition: the single-route model (e.g., Bybee & Moder, 1983), the dual-route model (e.g., Prasada & Pinker, 1993), and the multiple-rules model (e.g., Albright & Hayes, 2003).

Keywords: language acquisition, past tense, single-route model, dual-route model, multiple-rules model

A debate that lies at the heart of developmental psychology is whether children's linguistic development is best characterized in terms of abstract symbolic rules or processes such as analogy and abstraction over stored exemplars. A particular test case for this debate has been children's acquisition of the English past tense, a topic on which over 200 papers have been published (Pinker, 1999).

One reason why this domain has attracted so much attention is that the English past-tense system (unlike that of many morphologically richer languages) seems to exemplify a clear distinction between regular forms that are generated by a formal rule (add *–ed* to the stem, as in *walk/walked*) and irregular forms that are learned by rote or generated on phonological analogy with similar-sounding "neighbors" (e.g., *throw/threw, blow/blew, know/knew*). Children have long been known to be productive with the regular pattern, adding *–ed* to both irregular stems (e.g., *sing* → **singed*; asterisk indicates an ungrammatical or incorrect form) and novel stems presented experimentally (e.g., *wug* → *wugged*; Berko, 1958). The debate, then, is whether this productivity requires the notion of a formal context-free rule, as argued by proponents of the *dual-route* model (e.g., Prasada & Pinker, 1993), or whether it can be explained by a single analogical process operating on stored irregular and regular forms, as argued by proponents of the *single-route* model (e.g., Bybee & Moder, 1983). Recently, a third possibility has emerged. Under the *multiple-rules* model (Albright & Hayes, 2003) children set up an individual rule for each particular phonological context (i.e., each stem → past-tense mapping type), whether notionally irregular (e.g., *–ing* → *–ang*) or regular (*–sh* → *–shed*). The goal of this article is to test the predictions of these competing models, using a paradigm that is entirely novel to this domain: child grammaticality judgments of novel forms.

According to the dual-route account (e.g., Pinker & Prince, 1988; Pinker & Ullman, 2002; Prasada & Pinker, 1993) irregular and regular past-tense forms are produced by different mechanisms (via different routes). Irregular forms are retrieved from memory, or generated by analogy with irregular phonological neighbors stored in an associative memory system. Regular forms may sometimes be retrieved from memory but can always be generated by the application of an abstract rule that (roughly speaking) adds *–ed* to the verb stem. When a past-tense form (e.g., for *throw*) is required, the system first checks the verb's lexical entry for a stored irregular form. If one is retrieved (e.g., *threw*), it will be produced. If not, the system then attempts to generate an irregular form by analogy with other irregulars in the same phonological neighborhood (also termed "family," "island of reliability," or group of "friends"; e.g., *blow/blew, know/knew*). If a form is generated with sufficient strength (i.e., there are enough sufficiently activated neighbors to support the analogy), then this form is output (e.g., *threw*). When an irregular form is retrieved or generated, this blocks output of the default rule. When this is not the case, the stem is inflected by the default rule. For *throw*, this process will yield an overregularization error (e.g., **throwed*), but for regular forms (e.g., *walk*), the correct form will be generated (e.g., *walked*).

It is important to note that this account does not explicitly prohibit storage of regular forms; the claim is simply that regular

This article was published Online First August 23, 2010.

This research was supported by Grant RES-062-23-0931 from the Economic and Social Research Council. Thanks are due to the schools and children who participated in this research; to Rachel Hadzik, Victoria Byrne, and Victoria Clark for data collection; to Danielle Matthews for help with the statistical analysis; and to Julian Pine for comments on a previous version of this article.

Correspondence concerning this article should be addressed to Ben Ambridge, School of Psychology, University of Liverpool, Eleanor Rathbone Building, Bedford Street South, Liverpool L69 7ZA, England. E-mail: Ben.Ambridge@Liverpool.ac.uk

forms do not have to be stored (e.g., Pinker & Ullman, 2002). For example, *walked* could be generated by application of the default rule to the stem *walk* or retrieved directly from memory. This leaves open the possibility that some versions of the dual-route model could allow regular past-tense forms to be generated by analogy to stored regular forms—a point to which we return in the discussion. However, the present study (like those of Prasada & Pinker, 1993, and Albright & Hayes, 2003, upon which it is based) tests a strong version of the dual-route model, under which novel forms may be generated by phonological analogy to stored irregulars but not regulars. As the comparison between the dual-route model and its rivals presented in this article hinges crucially on this assumption, it is worth quoting the leading advocates of the dual-route model directly on this point. Discussing the findings of their own novel-verb judgment study (with adults), Prasada and Pinker (1993) note that "the goodness of the suffixed [i.e., regular] past-tense forms does not decline as a function of distance from known suffixed forms" (p. 22) and take this finding as support for the dual-route model over rival accounts. Accordingly, the present study also tests the dual-route model by investigating whether the acceptability of novel regular forms varies as a function of similarity to existing regular forms.

Under the single-route model (e.g., Bybee & Moder, 1983), all non-rote-produced past-tense forms, whether notionally regular or irregular, are generated by phonological analogy to stored regular and irregular forms by means of an associative memory system (similar to that proposed for irregulars under the dual-route model). For example, if a child fails to retrieve *threw*, the system may generate either *threw*—by analogy with *blow/blew, know/knew* (as under the dual-route model)—or **throwed*, by analogy with *show/showed, glow/glowed* (and, in a more distributed fashion, by analogy with all past-tense forms ending in *–d*).

Which of these two possible outcomes occurs depends primarily on the relative number of verbs exemplifying the *blow/blew* and *show/showed* patterns stored in memory (i.e., the type frequency of the pattern). Other factors such as the token frequency of individual verbs that undergo each pattern and the relative recent activation levels of the patterns (i.e., priming/fluency effects) may also influence the outcome.

Under the multiple-rules model (Albright & Hayes, 2003), a separate rule is set up for each particular stem→past-tense form mapping type (or phonological context), regardless of whether this would constitute a regular or irregular mapping under traditional accounts. For example, generalization over stem/past-tense mappings such as *bleed/bled* and *breed/bred* will yield a rule that transforms –[r/l]*eed* into –[r/l]*ed*, whereas generalization over mappings such as *miss/missed* and *kiss/kissed* will yield a rule that transforms *–iss* into *–issed*. In addition to storing rules that require a highly restrictive phonological context (as in the examples above), the learner iteratively collapses across these rules to form rules that are increasingly general (including an overgeneral rule that adds *–ed* to any verb stem). Rules at every level of generality are retained and assigned a confidence value based upon (a) the proportion of verbs for which the rule yields the correct past-tense form and (b) the raw number of forms used as the basis for forming the rule (as rules formed on the basis of more exemplars are more reliable). One consequence of this is that the more restrictive rules that specify a phonological context for a particular regular mapping (e.g., *–iss → –issed*) will be assigned a higher confidence value than the more general rule that adds *–ed* to any verb stem, as the latter will yield an incorrect form for a large proportion of verbs. When a past-tense form cannot be retrieved directly from memory, the rule that has the highest confidence value is used to generate the output.

Unlike the dual-route model, both the single-route and multiple-rules models predict that the acceptability of novel regular forms will increase as a function of similarity to existing regulars (though, as we will see in the Discussion, it may still be possible to choose between the two accounts). This is because the greater the similarity between a novel regular form and existing regulars, the greater the support for the relevant phonological analogy (single-route model) or the higher the confidence value of the relevant phonological rule (multiple-rules model). As we have already seen, the dual-route route model predicts no relationship between the acceptability of a novel regular form and its similarity to existing regulars (all three models predict similarity effects for irregulars).

A number of elicited production studies with children have attempted to test these competing predictions using real English verbs, sometimes with contradictory results. For example, whereas Marchman, Wulfeck, and Weismer (1999) found that overregularization errors were more common for irregular verbs that are similar to regulars (e.g., *throw*, similar to *show/showed*) than those that are not (e.g., *run*), an earlier study by Marchman (1997) found no such effect. This illustrates one problem with the use of real English verbs: Because there are only a finite number, it is difficult to select a stimulus set in which the number of regular and irregular "friends" and "enemies" can be varied systematically (friends/enemies are similar sounding verbs that form the past tense in the same/a different way; for example a friend of *sing/sang* is *ring/rang*, whereas an enemy is *bring/brought*). The problem is compounded by the necessity of matching verbs on other criteria such as frequency. Another obvious problem is that, in studies with real verbs, many past-tense forms will likely have been produced by rote.

Studies with adults have therefore generally used novel verbs. Again, however, the findings are somewhat contradictory. Whereas Prasada and Pinker (1993) found effects of similarity to regulars in a judgment task (Experiments 1 and 2), these effects disappeared when controlling for judgments of the acceptability of the stem forms. That is, novel regular forms that are not similar to existing regular families (e.g., *ploamphed*) were rated as less acceptable than novel forms that are similar to existing regular families (e.g., *plipped*). However, according to the authors, this was only because (as demonstrated in another study) participants considered *ploamph* to be less acceptable than *plip* as a stem form. Using a more carefully controlled set of stimuli, Albright and Hayes (2003) found an effect of similarity to existing regulars even when verbs were matched for stem acceptability.

The aim of the present study was to extend this paradigm—judgments of novel forms—to the population of primary interest: children who are still learning the English past-tense system. Children ages 6–7 years ($N = 20$) and 9–10 years ($N = 20$) rated the acceptability of the regular and irregular past-tense forms of 40 novel verbs, chosen to vary independently along the dimensions of similarity to existing regulars and similarity to existing irregulars.

These ages were chosen to allow for the investigation of developmental predictions of the three models, an issue that has been

rather neglected in this area (particularly in judgment studies, which hitherto have been conducted exclusively on adults). With regard to regular forms, the single-route and multiple-rules models assume that the older children will have a near-adultlike repertoire of the relevant families and hence show an effect of similarity to existing regulars on judgments of novel regulars. However, because most early acquired past-tense forms are irregular (see references in Rumelhart & McClelland, 1986), the younger group are unlikely to have an extensive repertoire of regular families. Thus these models predict an interaction whereby the effect of similarity to existing regulars on judgments of novel regulars is predicted to be significantly smaller for the younger than the older group (and, indeed, perhaps even nonexistent for the former). The dual-route model predicts no such developmental effect, as it predicts no effect of similarity to regulars at any age.

All models predict that children will show an effect of similarity to existing irregulars on judgments of novel irregulars. On the assumption that even the younger group will have acquired a near-adultlike repertoire of irregular families (because the majority of early acquired forms are irregular), no interaction with age is predicted.

Method

Participants

Participants were 40 normally developing monolingual learners of British English; 20 were ages 6 years 4 months to 7 years 2 months (M = 6 years 10 months), and 20 were ages 9 years 9 months to 10 years 8 months (M = 10 years 4 months), with an equal number of girls and boys at each age. All children were White and from a middle-class background (though detailed socioeconomic status [SES] information was not collected). Children were recruited via their school (in North West England); parents were sent a letter outlining the study and a consent form.

Design

The study used a correlational design across 40 items (novel verbs). The outcome variables were (a) acceptability ratings for the regular past-tense forms and (b) acceptability ratings for the irregular past-tense forms (with a separate set of regression analyses conducted for each). The predictor variables were (a) similarity of each novel verb to existing regular verbs, (b) similarity of each novel verb to existing irregular verbs, and (c) stem well-formedness (see Materials for details of how these were assessed). These variables, along with age group and gender, were included as fixed effects in the statistical analysis, with participant and item (verb) included as random effects.

The 40 verbs were split into two sets (A and B); at each age, half of the participants rated a regular and an irregular form of each verb in Set A and half rated a regular and an irregular form of each verb in Set B. This was necessary, as it was felt that young children would be unable to complete a full set of 80 trials. Preliminary mixed-effects regression models confirmed that on no occasion was set (included as a fixed effect) a significant predictor for either outcome measure ($t < 1$, $p = ns$ in all cases). Thus one can be confident that splitting the verbs between participants in this way had no significant effect on the outcome.

Materials

The 40 novel verbs chosen constituted the core set in Albright and Hayes' (2003) adult study. Where more than one irregular form is possible (e.g., *spling→splung/splang*), the form rated as most acceptable by adults was chosen for inclusion. The novel verbs were not generated by simply modifying existing regular and irregular forms (as in the study of Prasada & Pinker, 1993) but rather by using a computational procedure designed to ensure that (a) all stem forms are phonologically well-formed (e.g., avoiding forms such as *ploamph*) and (b) there exists a quantitative measure of precisely how similar each novel verb is to an existing class of regulars and an existing class of irregulars.

Specifically, Albright and Hayes (2003) constructed a set of 2,344 candidate forms "by concatenating combinations of relatively common syllable onsets and syllable rhymes" (p. 1353), which were then submitted to a computer instantiation of these authors' multiple-rules model. Using rules generated on the basis of the 4,253 stem/past-tense pairs from the English portion of the CELEX database (4,035 regular, 218 irregular), the model then produced an acceptability rating for the regular and irregular form of each candidate stem. These ratings constitute a measure of the similarity of each verb to a class of existing regulars and a class of existing irregulars (and also a prediction of the relative acceptability of each form by human raters). Using this metric, 40 verbs that vary continuously and orthogonally along the dimensions of similarity to regulars and similarity to irregulars were selected. Some example verbs are shown in Table 1 (for a full list, see Appendix A). It is important to emphasize that whereas, descriptively speaking, each verb can be characterized as similar to regulars only, irregulars only, both or neither, verbs in fact vary continuously along the dimensions of similarity to regulars and similarity to irregulars, both across and within these descriptive classes. Albright and Hayes's (2003) adult participants rated all 40 verbs for stem well-formedness using a 7-point scale, with a mean rating of 4.70 (SD = 0.69) indicating that participants generally considered them to be well formed.

Note that many of the 4,253 stem/past-tense pairs on the basis of which these stimuli were designed are likely to be unfamiliar to the younger group. This allows us to test the developmental predic-

Table 1
Example Stimuli

Similarity	Example
Similar to both an existing class of regulars and an existing class of irregulars	*dize (dized/doze), fro (froed/frew), rife (rifed/rofe)*
Similar to an existing class of regulars only	*bredge (bredged/broge), gezz (gezzed/gozz), nace (naced/noce)*
Similar to an existing class of irregulars only	*fleep (fleeped/flept), gleed (gleeded/gled), spling (splinged/splung)*
Not similar to either an existing class of regulars or an existing class of irregulars	*gude (guded/gude), nung (nunged/nang), preak (preaked/proke)*

Note. See Appendix A for a full list. Children's acceptability ratings were obtained for the regular and one possible irregular form of each verb.

tions outlined above, which are based on the assumption that the younger children have yet to encounter many (particularly regular) past-tense forms.

To make the study more engaging for children, an animation was created to illustrate each of the novel verbs. Each action was performed by a single character (a bunny, duck, frog, or bear) against a plain background and involved an unusual novel motion (e.g., dancing, squatting, bending arms and/or legs, etc.). The assignment of actions to novel verbs was randomized for each child (to control for possible effects of verb semantics on the choice of regular versus irregular inflection; see Ramscar, 2002). Animations were displayed on a laptop computer using Apple QuickTime Player. Children indicated their responses using a 5-point "smiley face" scale (see Figure 1) originally developed for obtaining judgments of argument structure overgeneralization errors, such as, **The joke giggled him* (see Ambridge, Pine, Rowland, Jones, & Clark, 2009; Ambridge, Pine, Rowland, & Young, 2008). An extensive outline and discussion of the strengths and weaknesses of this paradigm is given in Ambridge (in press).

Procedure

Children first completed a warm-up introducing them to the appropriate use of the scale. This warm up used regular and irregular noun plurals in both correct and incorrect form (*mouse/mice*, *house/*hice*, *foot/feet, book/*beek, box/boxes, man/*mans, eye/eyes, child/*childs, tooth/*tooths* and *mouse/*mouses*). Full details of the warm-up procedure can be found in Appendix B.

Children then completed the 40 test trials in pseudorandom order, with the constraint that the regular and irregular form of a given verb were always separated by at least four trials. For each trial, the experimenter presented the novel verb in the frame, "The [bunny/duck/frog/bear] likes to [VERB]. Look, there he is [VERB]ing. Every day he [VERB]s. So yesterday he [X]," where X is the regular or irregular past-tense form of the verb.

Results

For each of the two outcome measures—acceptability ratings for (a) novel regular past-tense forms and (b) novel irregular past tense forms—mixed-effects linear regression models (see Baayen, 2008, for an introduction) were fitted to the data (mean acceptability ratings for each regular and irregular past-tense form can be found in Appendix A). In addition to the variables of primary interest, all models included participant and item (verb) as random effects and gender and stem well-formedness (as determined by adult raters in Albright & Hayes, 2003) as fixed effects.

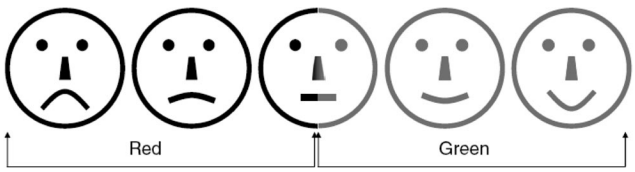

Figure 1. The scale used by children to rate the acceptability of individual past-tense forms. Reproduced from "The Effect of Verb Semantic Class and Verb Frequency (Entrenchment) on Children's and Adults' Graded Judgments of Argument-Structure Overgeneralization Errors," by B. Ambridge, J. M. Pine, C. F. Rowland, and C. R. Young, 2008, *Cognition, 106,* p. 105. © 2008 by Elsevier.

The critical predictions of the competing models relate to the effect of similarity to existing regulars on judgments of novel regulars. The dual-route model predicts no such effect at any age (i.e., that "the goodness of the suffixed [i.e., regular] past-tense forms does not decline as a function of distance from known suffixed forms"; Prasada & Pinker, 1993, p. 22). Both the single-route and multiple-rules models predict that the acceptability of novel regulars will increase as a function of similarity to existing regulars and that the magnitude of this effect will increase with age as more regular families are acquired.

The first analysis therefore investigated the effect of similarity to existing regulars (included as a fixed effect) on children's ratings of the acceptability of novel regular forms. Also included as fixed effects were (a) children's ratings of the corresponding novel irregular form (to control for any trade-off effects occurring during the judgment task); (b) similarity to existing irregulars (to check that any apparent effect of similarity to existing regulars was not, in fact, an effect of dissimilarity to existing irregulars); and (c) age group and Age Group $\times$ Similarity to Regulars (to investigate any possible developmental changes in the effect of similarity to regulars).

This model is summarized in Table 2. As predicted by the single-route and multiple-rules models, a significant effect of similarity to existing regulars was observed (as well as a trade-off effect). However, this effect must be interpreted in the light of a significant interaction of similarity to existing regulars by age group. This interaction was investigated by fitting a separate regression model to the data for each age group (see Table 2). This analysis revealed that the source of the interaction was the developmental effect predicted by the single-route and multiple-rules models: The effect of similarity to regulars was larger for the older group ($\beta = 6.70$, $SE = 1.88$, $t = 3.56$, $p < .001$) than the younger group ($\beta = 1.65$, $SE = 1.52$, $t = 1.08$, $p = .28$, *ns*) and indeed did not reach significance for the younger group independently.

The second set of analyses investigated the effect of similarity to existing irregulars (included as a fixed effect) on children's ratings of the acceptability of novel irregular forms, with the same controls (i.e., ratings of the corresponding novel regular form and similarity to existing regulars) in place. All three models predict a significant effect of similarity to existing irregulars, and no—or a very small—increase in the magnitude of this effect with age. The model fitted to these data is summarized in Table 3. As predicted, a significant effect of similarity to irregulars was observed (as well as a trade-off effect), with no similarity by age group interaction.

Discussion

In the present study, children rated regular and irregular past-tense forms of 40 novel verbs, chosen to vary independently along the dimensions of similarity to existing regular forms and similarity to existing irregular forms. As predicted by all three models under investigation, the acceptability of novel irregulars was shown to increase as a function of similarity to existing irregulars, with no evidence to suggest an increase in the magnitude of this effect with age.

As predicted by the single-route and multiple-rules model, but not the dual-route model, the acceptability of novel regulars was

Table 2
Fixed Effects for Regression Model Fitted to Acceptability Ratings of Novel Regular Past-Tense Forms

Variable	β	HPD intervals		*SE*	*t*	*p*
		Lower	Upper			
All children[a]						
(Intercept)	14.02	4.87	23.61	4.72	2.97	0.00
Gender	0.18	−0.23	0.62	0.30	0.62	0.54
Stem well-formedness	0.04	−0.07	0.15	0.06	0.67	0.50
Rating for irregular	**−0.08**	**−0.14**	**−0.02**	**0.03**	**−2.57**	**0.01**
Similarity to irregulars	0.20	−0.04	0.46	0.13	1.60	0.11
Similarity to regulars	**−10.46**	**−19.87**	**−0.62**	**4.82**	**−2.17**	**0.03**
Age group	**−1.91**	**−3.11**	**−0.69**	**0.61**	**−3.12**	**0.00**
Similarity to Regulars × Age Group	**1.95**	**0.73**	**3.23**	**0.63**	**3.12**	**0.00**
6- to 7-year-olds[b]						
(Intercept)	1.84	−1.23	4.87	1.52	1.21	0.23
Gender	0.33	−0.30	0.96	0.43	0.77	0.44
Stem well-formedness	0.09	−0.08	0.22	0.07	1.15	0.25
Rating for irregular	−0.07	−0.14	0.02	0.04	−1.64	0.10
Similarity to irregulars	0.32	−0.01	0.67	0.17	1.88	0.06
Similarity to regulars	1.65	−1.34	4.81	1.52	1.08	0.28
9- to 10-year-olds[c]						
(Intercept)	−2.43	−5.96	0.94	1.87	−1.30	0.19
Gender	0.04	−0.56	0.67	0.42	0.09	0.93
Stem well-formedness	−0.01	−0.16	0.18	0.09	−0.07	0.94
Rating for irregular	−0.08	−0.17	0.01	0.05	−1.76	0.08
Similarity to irregulars	0.08	−0.33	0.45	0.21	0.40	0.69
Similarity to regulars	**6.70**	**3.16**	**10.17**	**1.88**	**3.56**	**0.00**

Note. Bold values indicate that effect is statistically significant at $p < .05$ or greater. HPD = highest probability density.
[a] Model log likelihood = −1,166. Random effects: Participant (Var = 0.84, *SD* = 0.92), Verb (Var = 0.01, *SD* = 0.08). [b] Model log likelihood = −582.7. Random effects: Participant (Var = 0.89, *SD* = 0.94), Verb (Var < 0.001, *SD* < 0.001). [c] Model log likelihood = −581.7. Random effects: Participant (Var = 0.82, *SD* = 0.91), Verb (Var = 0.05, *SD* = 0.23).

shown to increase as a function of similarity to existing regulars, with the magnitude of this effect increasing with age. Indeed, when the data were analyzed separately, the effect of similarity to regulars did not reach significance for the younger group, as would be expected on the assumption that these children have yet to form many of the regular families (because most early acquired past-tense forms are irregular).

Before considering whether it is possible to posit a version of the dual-route model that can account for these findings, it is important to ask if there is any way that they can be used to choose between the single-route and multiple-rules models, both of which predict the observed effects. In principle, one could compare the two models by instantiating each as a computer simulation and taking each model's output for the regular and irregular form of each verb as its prediction of the relevant human acceptability judgment score. The model that best simulates the pattern of human judgments (at each age) would then be declared the winner. For adult data, Albright and Hayes (2003) found that when the two models are compared in this way, the multiple-rules model outperforms the single-route model.

However, we do not attempt to compare the models on their ability to predict the pattern of child judgments, as subsequent

Table 3
Fixed Effects for Regression Model Fitted to Acceptability Ratings of Novel Irregular Past-Tense Forms

Variable	β	HPD intervals		*SE*	*t*	*p*
		Lower	Upper			
(Intercept)	−2.63	−14.01	7.83	5.59	−0.47	0.64
Gender	0.14	−0.25	0.53	0.24	0.58	0.56
Stem well-formedness	0.11	−0.06	0.29	0.09	1.20	0.23
Rating for regular	**−0.10**	**−0.17**	**−0.01**	**0.04**	**−2.40**	**0.02**
Similarity to irregulars	**0.69**	**0.28**	**1.08**	**0.21**	**3.24**	**0.00**
Similarity to regulars	5.21	−6.18	16.35	5.74	0.91	0.36
Age group	0.88	−0.44	2.33	0.70	1.25	0.21
Similarity to Irregulars × Age Group	−0.95	−2.44	0.44	0.72	−1.31	0.19

Note. Bold values indicate that effect is statistically significant at $p < .05$ or greater. Model log likelihood = −1,277. Random effects: Participant (Var = 0.50, *SD* = 0.71), Verb (Var = 0.08, *SD* = 0.30). HPD = highest probability density.

simulations of the adult data have shown that the outcome depends crucially on the way that the single-route model is implemented computationally. The single-route model tested by Albright and Hayes (2003) was an *exponential-decay* single-route model under which, when determining the past-tense form of a novel verb (e.g., *spling*), the nearest neighbor (e.g., *spring* → *sprung*) has the greatest influence; the influence of other neighbors (e.g., *ring* → *rang*) decreases exponentially as phonological distance increases. Although the exponential-decay single-route model was outperformed by the multiple-rules model, Keuleers and Daelemans (2007; see also Keuleers, 2008) demonstrated that both were outperformed by a $k = 7$ zero-decay single-route model, under which each neighbor exerts equal influence (with the number of neighbors [k] specified in advance as a model parameter).

At this stage, then, it is probably wise to reserve judgment with regard to whether the single-route or multiple-rules model offers the best account of the data. Future computer modeling studies, using the present child judgment data as the target, would be valuable in addressing this issue.

The version of the dual-route model that has been considered up to this point is the traditional dual route model as outlined, for example, in Prasada and Pinker (1993), and tested in both Prasada and Pinker (1993) and Albright and Hayes (2003). Recall that Prasada and Pinker (1993) took as support for the dual-route model their finding that "the goodness of the suffixed past-tense forms does not decline as a function of distance from known suffixed forms" (p. 22). Clearly, this prediction follows only from a version of the dual-route model under which analogy to stored regular forms plays no (or very little) role. Equally clearly, the present finding that the goodness of suffixed (regular) past-tense forms does decline as a function of distance from known suffixed (regular) forms counts against the Prasada and Pinker (1993) version of the dual-route model.

Of course, this finding does not count against a version of the dual-route model that assumes that many (or perhaps all) regular past-tense forms are also stored in memory and used as the basis for analogical generalization. The model of Alegre and Gordon (1999) is sometimes cited as an example of such an account, on the basis that it assumes storage of regulars that meet a certain frequency threshold. In fact, however, these authors are clear in their claim that "associative mechanisms arise only in the irregular vocabulary" (Alegre and Gordon, 1999, p. 57). One dual-route model that explicitly does assume analogy on the basis of stored regulars is that of Hartshorne and Ullman (2006; for evidence against this account, see Kidd & Lum, 2008). Although this version of the dual-route model can account for the present data, it is not easy to see how one could test this model against the rival accounts empirically.

One possibility, however, is as follows. Presumably if at least some novel forms are not generated by analogy to stored regulars but by a default rule, one would expect to see a smaller effect of similarity to regulars than similarity to irregulars, because all novel irregulars are generated by analogy to stored irregulars. Although the present data are consistent with this possibility (the t values for similarity to irregulars and similarity to regulars are 3.24 and 2.17, respectively), this is not the case for adults: Albright and Hayes (2003) report partial r values of 0.49 and 0.58, respectively (recall also that the single-route and multiple-rules models predict that the effect of similarity to regulars will be small or nonexistent early in development, when few regular families have been learned). The finding that, for adults, the effect of similarity to regulars is equal to (or even greater than) the effect of similarity to irregulars can be reconciled only with a version of the dual-route model that assumes that (virtually) all regulars are stored and used as the basis for analogical generalization. This is not to say that the default rule would play no role. For example, it may be required for inflecting denominal regular verbs that are homophonous with regulars, as in *The army ringed the city* (Kim, Pinker, Prince, & Prasada, 1991; Marcus et al., 1992; but see Hahn & Nakisa, 2000; Plunkett & Juola, 1999; Ramscar, 2002), an issue that is not addressed in the present study.

To conclude, the finding of a correlation between similarity to existing regulars and ratings for novel regular forms observed for 9- to 10-year-olds in the present study (as well as for the adults studied by Albright & Hayes, 2003) means that a successful model of past-tense learning will have to incorporate a role for either (a) analogy operating over stored regular forms or (b) multiple regular rules formed on the basis of encountering such forms in the input. Clearly, the single-route and multiple-rules models meet this criterion, as do dual-route models that—in addition to the default regular rule—assume analogical generalization over stored regulars. Any successful model must also account for the observed developmental pattern whereby the effect of similarity to regulars is small (indeed, nonexistent) for young children but becomes at least as large as the effect of similarity to irregulars by adulthood (presumably as more regular forms are learned, or the regular rules strengthened). Future studies, using both computer modeling and experimental paradigms, should attempt to further tease apart the predictions of the models, to move the field closer to an understanding of whether past-tense inflection (and, by extension, language learning in general) is governed by formal rules or analogy to stored exemplars.

References

Albright, A., & Hayes, B. (2003). Rules vs. analogy in English past tenses: A computational/experimental study. *Cognition, 90,* 119–161.

Alegre, M., & Gordon, P. (1999). Frequency effects and the representational status of regular inflections. *Journal of Memory and Language, 40,* 41–61.

Ambridge, B., Pine, J. M., Rowland, C. F., Jones, R. L., & Clark, V. (2009). A semantics-based approach to the "no negative-evidence" problem. *Cognitive Science, 33,* 1301–1316.

Ambridge, B., Pine, J. M., Rowland, C. F., & Young, C. R. (2008). The effect of verb semantic class and verb frequency (entrenchment) on children's and adults' graded judgments of argument-structure overgeneralization errors. *Cognition, 106,* 87–129.

Ambridge, B. (in press). Assessing grammatical knowledge. In E. Hoff (Ed.), *The guide to research methods in child language.* New York, NY: Wiley-Blackwell.

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R.* Cambridge, England: Cambridge University Press.

Berko, J. (1958). The child's learning of English morphology. *Word, 14,* 150–177.

Bybee, J. L., & Moder, C. L. (1983). Morphological classes as natural categories. *Language, 59,* 251–270.

Hahn, U., & Nakisa. (2000). German inflection: Single route or dual route. *Cognitive Psychology, 47,* 313–360.

Hartshorne, J. K., & Ullman, M. T. (2006). Why girls say "holded" more than boys. *Developmental Science, 9,* 21–32.

Keuleers, E. (2008). *Memory-based learning of inflectional morphology.* (Unpublished doctoral dissertation). University of Antwerp, Antwerp, Belgium.

Keuleers, E., & Daelemans, W. (2007). Memory-based learning models of inflectional morphology: A methodological case study. *Lingue E Linguaggio, 6*(2), 151–174.

Kidd, E., & Lum, J. (2008). Sex differences in past tense overregularization. *Developmental Science, 11,* 882–889.

Kim, J. J., Pinker, S., Prince, A., & Prasada, S. (1991). Why no mere mortal has ever flown out to center field. *Cognitive Science, 15,* 173–218.

Marchman, V. A. (1997). Children's productivity in the English past tense: The role of frequency, phonology and neighborhood structure. *Cognitive Science, 21,* 283–304.

Marchman, V. A., Wulfeck, B., & Weismer, S. E. (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.

Marcus, G. F., Pinker, S., Ullman, M., Hollander, M., Rosen, T. J., & Xu, F. (1992). Overregularization in language acquisition [Monograph]. *Monographs of the Society for Research in Child Development, 57*(54, Serial No. 228).

Pinker, S. (1999). *Words and rules: The ingredients of language.* New York, NY: Basic Books.

Pinker, S., & Prince, A. (1988). On language and connectionism: Analysis of a parallel distributed processing model of language acquisition. *Cognition, 28,* 1–2.

Pinker, S., & Ullman, M. T. (2002). The past and future of the past tense. *Trends in Cognitive Sciences, 6,* 456–463.

Plunkett, K., & Juola, P. (1999). A connectionist model of English past tense and plural morphology. *Cognitive Science, 23,* 463–490.

Prasada, S., & Pinker, S. (1993). Generalisation of regular and irregular morphological patterns. *Language and Cognitive Processes, 8,* 1–56.

Ramscar, M. (2002). The role of meaning in inflection: Why the past tense does not require a rule. *Cognitive Psychology, 45,* 45–94.

Rumelhart, D. E., & McClelland, J. L. (1986). *Parallel distributed processing: Explorations in the microstructure of cognition.* Cambridge, MA: MIT Press.

Appendix A

Stimulus Set for the Present Study, With Mean Ratings From Child (Present Study) and Adult (Albright & Hayes, 2003) Participants

	Verb forms					Acceptability judgments						Model prediction (similarity to)	
						Age 6–7		Age 9–10		Adults			
No.	Verb	Present	Regular past	Irregular past	Well formed	Reg.	Irreg.	Reg.	Irreg.	Reg.	Irreg.	Reg.	Irreg.
1	bize	bÈYz	bÈYzd	bÈoz	4.57	3.90	2.00	3.80	3.10	5.30	4.57	0.99	0.12
2	dize	dÈYz	dÈYzd	dÈoz	4.62	3.70	3.40	4.40	2.40	5.42	5.04	0.99	0.55
3	drice	drÈYs	drÈYst	drÈos	3.86	4.40	2.10	3.30	2.70	6.26	4.48	1.00	0.85
4	flidge	flÈIJ	flÈIJd	flÈÃJ	4.05	3.70	2.30	4.40	2.30	6.21	4.88	0.99	0.20
5	fro	frÈo	frÈod	frÈu	5.84	4.80	3.10	3.30	4.70	5.83	4.33	0.94	0.72
6	gare	gÈer	gÈerd	gÈor	5.24	3.40	2.70	4.10	2.30	6.57	3.39	0.98	0.26
7	glip	glÈIp	glÈIpt	glÈÃp	4.95	4.30	2.20	3.70	2.50	5.95	3.45	0.99	0.06
8	rife	rÈYf	rÈYft	rÈof	5.61	3.70	2.60	4.50	1.90	5.95	4.14	1.00	0.47
9	stin	stÈIn	stÈInd	stÈÃn	5.40	4.00	2.60	3.50	2.50	5.30	4.78	0.97	0.28
10	stip	stÈIp	stÈIpt	stÈÃp	5.45	4.00	3.00	4.60	1.90	5.92	4.50	0.99	0.15
11	blig	blÈIg	blÈIgd	blÈÃg	3.71	4.40	2.40	3.80	2.80	5.67	4.17	0.96	0.88
12	chake	CÈek	CÈekt	CÈUk	5.33	3.70	3.60	3.90	1.70	5.74	5.04	0.90	0.83
13	drit	drÈIt	drÈIt«d	drÈIt	4.30	3.50	4.10	3.90	2.60	4.96	5.13	0.94	0.48
14	fleep	flÈip	flÈipt	flÈEpt	4.24	3.70	3.60	4.30	3.40	5.00	6.09	0.96	0.85
15	gleed	glÈid	glÈid«d	glÈEd	5.29	4.30	2.40	2.90	3.70	4.22	6.00	0.87	0.79
16	glit	glÈIt	glÈIt«d	glÈIt	5.25	3.70	3.50	3.90	3.30	5.00	5.21	0.94	0.67
17	queed	kwÈid	kwÈid«d	kwÈEd	3.81	3.40	2.80	4.10	1.90	4.65	5.35	0.87	0.35
18	plim	plÈIm	plÈImd	plÈÃm	4.43	4.30	1.90	3.30	2.70	6.13	4.17	0.97	0.41
19	skride	skrÈYd	skrÈYd«d	skrÈod	4.05	3.90	2.20	2.80	3.00	4.17	4.39	0.89	0.73
20	spling	splÈIN	splÈINd	splÈÃN	4.05	3.90	2.60	3.50	3.50	4.36	5.45	0.92	0.88
21	teep	tÈip	tÈipt	tÈEpt	4.95	4.20	2.60	3.60	3.00	5.91	4.70	0.96	0.56
22	gude	gÈud	gÈud«d	gÈud	4.25	4.00	3.90	2.80	2.60	4.90	5.55	0.99	0.01
23	nung	nÈÃN	nÈÃNd	nÈQN	3.21	3.80	1.50	3.10	2.40	5.37	4.32	0.92	0.00
24	pank	pÈQNk	pÈQNkt	pÈÃNk	5.62	3.90	3.10	4.30	1.60	6.30	4.00	0.96	0.00
25	preak	prÈik	prÈikt	prÈok	4.91	4.20	1.90	2.10	3.20	5.83	3.92	0.94	0.03
26	rask	rÈQsk	rÈQskt	rÈÃsk	5.30	3.80	2.80	4.80	2.40	6.42	4.08	0.98	0.00

(*Appendices continue*)

Appendix A (*continued*)

	Verb forms					Acceptability judgments						Model prediction (similarity to)	
						Age 6–7		Age 9–10		Adults			
No.	Verb	Present	Regular past	Irregular past	Well formed	Reg.	Irreg.	Reg.	Irreg.	Reg.	Irreg.	Reg.	Irreg.
27	shilk	SÈIlk	SÈIlkt	SÈQlk	4.60	4.30	2.60	3.90	2.70	5.79	3.67	0.98	0.00
28	tark	tÈark	tÈarkt	tÈork	5.10	3.90	2.90	4.30	1.60	6.33	3.71	0.98	0.00
30	trisk	trÈIsk	trÈIskt	trÈQsk	5.14	3.80	2.40	3.90	1.80	6.29	3.76	0.98	0.00
31	blafe	blÈef	blÈeft	blÈEft	3.57	4.20	1.70	3.70	2.60	6.32	4.09	1.00	0.00
32	bredge	brÈEJ	brÈEJd	brÈoJ	3.86	3.50	2.70	4.20	1.80	6.33	3.43	0.99	0.00
33	chool	CÈul	CÈuld	CÈol	3.76	4.20	2.70	3.90	2.30	6.13	3.71	0.99	0.00
34	dape	dÈep	dÈept	dÈQpt	5.14	4.00	2.10	4.80	2.20	6.25	4.00	0.99	0.00
35	ghez	gÈEz	gÈEzd	gÈaz	4.19	3.90	2.10	4.20	2.20	6.61	2.52	0.99	0.00
36	nace	nÈes	nÈest	nÈos	5.00	3.60	2.00	4.60	1.50	6.57	2.91	1.00	0.05
37	spack	spÈQk	spÈQkt	spÈÃk	5.05	4.00	2.30	3.30	2.80	6.22	3.96	0.98	0.00
38	stire	stÈYr	stÈYrd	stÈor	5.62	3.00	2.70	4.10	2.10	6.00	3.22	0.98	0.00
39	tesh	tÈES	tÈESt	tÈaS	4.71	4.30	1.80	4.20	2.20	6.22	3.13	1.00	0.00
40	whiss	wÈIs	wÈIst	wÈÃs	5.76	3.50	2.50	4.20	1.60	6.57	3.35	1.00	0.00
41	nold	nÈold	nÈold«d	nÈold	4.10	3.50	3.10	3.70	3.10	4.64	6.05	0.90	0.01

Note. The unscaled predictions of the multiple-rules model (Albright & Hayes, 2003) constitute the measure of similarity to regulars/irregulars used in the present study. Data on adult participants were adapted from "Rules vs. Analogy in English Past Tenses: A Computational/Experimental Study," by A. Albright & B. Hayes, 2003, *Cognition, 90,* pp. 155–157. © 2003 by Elsevier.

Appendix B

Warm-Up Procedure

The experimenter introduced the procedure by saying, "I'm going to say some words. Sometimes I say them right, but sometimes I get it wrong and say them a bit funny. Can you tell me when I get it right and when I get it wrong? When I say it right, we're going to choose the green counter and put it here [places green counter on happiest face]. When I say it wrong, we're going to choose the red counter and put it here [places red counter on saddest face]. Don't worry about the other faces for now." The experimenter then completed the first two trials, using the relevant (still) pictures. For the maximally acceptable sentence, the experimenter said, "Here is one mouse. Now there are two mice." For the maximally unacceptable sentence, the experimenter said, "Here is one house. Now there are two hice" (i.e., an irregularization error). The experimenter then invited the child to complete the next two trials (*foot/feet* and *book/beek*), correcting the child if she did not choose Faces 5 (i.e., happiest) and 1 (i.e., saddest), respectively. The experimenter then continued, saying, "Now sometimes I say it right but it's not perfect. If it's good but not perfect, you can put the green counter here [Face 4]. If it's a little bit right and a little bit wrong, or somewhere in between you can put it here [Face 3]. Also, sometimes I say it wrong but it's not really terrible. If it's wrong but not terrible, you can put the red counter here [Face 2]. If it's a little bit wrong and a little bit right, or somewhere in between, you can put it here [Face 3]." The experimenter then invited the child to complete the remaining practice trials (*two boxes* [5]; *two mans* [1–2], *two eyes* [5]; *three childs* [1–2]; *two tooths* [2–3] and *three mouses* [2–3]), encouraging the child to select the responses indicated. The purpose of this warm-up was to introduce the child to the procedure, including the fact that both regular and irregular responses could be either unacceptable or acceptable, although the use of any past-tense forms was avoided.

Received April 8, 2009
Revision received June 23, 2010
Accepted June 25, 2010 ■