Mechanisms of Verb Inflection – Regular vs. Irregular or Easy vs. Hard?

Gert Westermann (gwestermann@brookes.ac.uk)
Vanja Kovic (vvucetic@brookes.ac.uk)
Nicolas Ruh (nruh@brookes.ac.uk)

Department of Psychology, Oxford Brookes University
Oxford OX3 0BP, UK

Abstract

We present a speeded production study of the English past tense that is designed to evaluate between conflicting theories of the mechanism of verb inflection. Test items are verbs that systematically differ in regularity but also in overall ease of processing as defined by a number of statistical and distributional factors. We show that overall ease, but not regularity, predicts differences in the response latencies of produced past tense forms in adults. These results indicate that the processing of inflections is best explained by a single, associative mechanism that is responsible for the production of all verbs.

Keywords: English past tense; verb inflection; speeded production task

Introduction

The mechanisms underlying the processing of the English past tense have during the past 20 years become a topic of intense investigation and controversial debate. Part of the reason for such a small aspect of language processing taking such a prominent role is that it touches on wider issues, namely, the mental reality of linguistic rules. The regular past tense is the classic case of such a rule: to form the past tense of a verb, add –ed to its stem (such as looked). However, not all verbs form their past tense in this way. There are also around 160 irregular verbs with idiosyncratic past tense forms (such as write-wrote and sing-sang). Two main theories have emerged from the past tense debate: the Dual Mechanism, or Words and Rules, Theory (DMT, e.g., Pinker, 1999; Pinker & Ullman, 2002) holds that regular and irregular forms are processed by two separate mechanisms: a symbolic rule by which regular forms are generated on the fly, and retrieval from an associative mental lexicon of irregular forms. Evidence for the DMT has been collected by describing apparent dissociations between regular and irregular forms across a range of studies. For example, elicitation studies in which participants are asked to produce the past tense form of a presented stem have shown that frequent irregular forms were named faster than infrequent ones, whereas no or less pronounced differences were found between frequent and infrequent regulars (Prasada et al., 1990; Albright & Hayes, 2003). In the acquisition of the past tense children produce many overregularization errors such as comed but very few irregularizations (Marcus et al., 1992). Agrammatic aphasic patients have been argued to display selective deficits in the production of regular past tense forms (Ullman et al., 1997), whereas patients with semantic dementia show selective deficits in processing irregular forms (Patterson et al., 2001). Further dissociations between regular and irregular forms have been revealed in brain imaging studies (e.g., Newman et al., 2002; Tyler et al., 2005).

However, these results have not gone unchallenged. A contrasting theory of verb inflection argues that all forms, regulars and irregulars alike, are processed in a single, associative mechanism (Rumelhart & McClelland, 1986; McClelland & Patterson, 2002). According to this view, which is closely linked to connectionist modeling, processing differences between verbs emerge because all verbs compete for the same resources in a single system, leading to effects of statistical factors such as frequency and phonological properties. This single mechanism view predicts that there is no qualitative, clear-cut dissociation between regulars and irregulars but instead a graded continuum of differences based on these statistical factors. Several empirical studies have backed this view by showing that apparent dissociations between regulars and irregulars can be explained by confounding statistical factors. For example, Marchman (1997) investigated the past tense errors made by school-age children and found that all errors were predicted by a combination of statistical factors such as frequency, phonological complexity of stem and past tense form, and the numbers of ‘friends’ (similar sounding stems with the same past tense such as looked and cooked) and ‘enemies’ (similar sounding stems with different past tense forms such as hear-heard and feared). In another study, Bird et al. (2003) found that differences in errors for regular and irregular inflection that had been claimed to exist in Broca’s aphasics (Ullman et al., 1997) disappeared when the verbs were matched for phonological complexity. Differences in the activation of brain areas for regular and irregular verbs (Beretta et al., 2003) were used by Seidenberg and Arnoldussen (2003) to argue that irregular verbs are harder to process and therefore activate more brain areas than regulars if distributional factors are not controlled for. In an fMRI study, Joanisse & Seidenberg (2005) found that apparent differences in brain activation between regular and irregular verbs were better explained by phonological characteristics of the past tense forms.
Proponents of the DMT have reacted to the evidence that regular verbs often show characteristics of associative storage by accepting that regular forms can, but need not be, stored together with irregulars (Pinkser, 1999). However, unless it is precisely specified under which circumstances a regular verb becomes stored this modification reduces the value of the DMT to a post-hoc descriptive theory: if a regular verb shows storage properties it must be stored, if it does not it may have been produced by the rule mechanism. It also makes it difficult to see how this theory could be falsified.

Whereas the single mechanism view offers a more parsimonious account of verb inflection, it has not yet been investigated in detail which factors lead to processing differences between verbs. Different studies have each focused on a number of such factors: phonological friends and enemies, frequency, and presence of a stem final alveolar (t/d) (Marchman, 1997), imageability and phonological complexity (Bird et al., 2003), similarity between base and past tense form (Rueckl et al., 1997) and concreteness, imageability and phonological similarity (Joanisse & Seidenberg, 2005). Several authors have made the more general point that combinations of these factors make verbs easier or harder to process, leading to observable processing differences (Westermann, 2000; Seidenberg & Arnoldussen, 2003). In this paper we take up this idea and develop a classification for verbs based on their overall ‘easiness’. This classification then allows us to directly evaluate single and dual mechanism accounts against each other by selecting verbs for experimental studies that vary both in their easiness and their regularity.

Here we argue that the factors underlying processing differences in a single system can be understood from a connectionist perspective. In a neural network model several factors affect how easy or hard it is to learn and process a certain word. In most models of language processing, the representations for different forms overlap, making it easier for similar inputs to map to similar outputs. Similar inputs requiring different outputs are harder to learn. Thus, phonological friends would make a verb easier, and phonological enemies, harder. The frequency of input-output mappings is also an important factor, with high frequency forms being easier as they will lead to more weight change in a connectionist system than less frequent forms. Age of acquisition is another possible easiness factor with early acquired forms having a stronger impact on the system than later forms (Ellis & Lambon-Ralph, 2000) and therefore being easier. Finally, previous empirical work suggests that presence of a final alveolar (t/d) in a verb stem makes it more difficult to inflect because the stem already appears to be an inflected form (Stemberger & MacWhinney, 1986).

Here we classify verbs as overall easy or hard based on these factors. We can then define easy and hard regulars and irregulars to directly evaluate the DMT and single mechanism approaches against each other. The DMT predicts dissociations between regular and irregular verbs irrespective of their easiness. Single mechanism approaches predict graded dissociations based on easiness, irrespective of whether verbs are regular or irregular. In order to test whether regularity or easiness lie at the basis of processing differences between verbs, we conducted speeded production experiments with children and adults.

Previous speeded production studies have yielded conflicting results. In most cases they had a 2x2 design in which low and high frequency regular and irregular forms were used. A study by Prasada et al. (1990) found response time differences between low and high frequency irregulars but no differences for regulars. Another study by Beck (1997) found the same result in one experiment and a longer response time for high frequency than for low frequency regulars in another experiment. Such an anti-frequency effect was also found in a study with German children and adults (Clahsen et al., 2004). From an easy-hard perspective these conflicting results could be argued to arise from the fact that only frequency but not other statistical factors were controlled for.

**Experiment**

To assess whether processing differences in past tense production are mainly affected by the grammatical category of a verb (regular or irregular) as predicted by the DMT or by a combination of statistical factors as predicted by single mechanism approaches, we developed a set of verbs in which these factors were systematically varied to obtain easy, intermediate and hard regular and easy, intermediate and hard irregular verbs. These verbs were then used in a speeded production task in which participants were auditorily presented with a verb stem and had to say out loud the corresponding past tense form as fast as possible. Differences in response latencies for different verbs could then be analyzed with respect to the two competing theories. To investigate developmental change in past tense processing we tested 10-11 year old children as well as adults.

**Method**

**Participants.** For the studies with children, 18 10-11 year olds (mean age: 11;0) were recruited from a school in Oxfordshire. Head teacher consent was obtained prior to testing. Furthermore, 18 adult participants were recruited from Oxford Brookes University and the University of Oxford. All participants were native speakers of British English and had normal hearing.

**Stimuli.** All monosyllabic verbs with a lemma frequency of greater than 200 in 17.9 million words were extracted from the English part of the CELEX database (Baayen et al., 1993). Each verb was then characterized according to the following features: log stem frequency, log past tense frequency, log stem and log past frequency in parental input (Theakston et al., 2001), stem final alveolar consonant, phonological complexity (defined as number of phonemes per syllable), number of ‘friends’ types and tokens, number
of ‘enemy’ types and tokens, ratio of friends to enemies (types and tokens). Where age of acquisition and imageability norms were available (Bird et al., 2001) they were also used. For phonological complexity, number of enemy types/tokens and age of acquisition, a verb was assigned a value of 1 (easy) if it fell below the median of this measure for all verbs and a value of 0 (hard) otherwise. For the other features a 1 (easy) was assigned if the value fell above the median and a 0 (hard) if it fell below the median. A total ‘easiness’ score was then computed for each verb as the mean of its scores for individual features. Thus, a verb scoring close to 1 was overall easy, and a verb scoring close to 0 was overall hard.

From this corpus a set of 60 verbs was extracted so that there were 30 regular and 30 irregular verbs. For the regular verbs there were 13 easy, 10 intermediate and 7 hard verbs, and for irregulars, there were 7 easy, 14 intermediate and 9 hard verbs. These verbs were digitally recorded by a female speaker with a sampling rate of 48kHz. The mean duration of a verb stem was as follows: easy regular: 710 ms, intermediate regular: 642 ms, hard regular: 713 ms, easy irregular: 627 ms, intermediate irregular: 707 ms, hard irregular: 758 ms. There was no significant difference between stem duration for regular and irregular (F(1) = 0.131, p=0.718) or easy, intermediate and hard (F(2) = 2.608, p=0.083) verbs.

**Design.** The two participant groups were tested on their ability to produce the past tense of 60 English verbs, differing on grammatical regularity (regular or irregular) and ease of production (easy, intermediate or hard), based on the statistical factors discussed above. Each participant processed the same 60 verbs, consisting of an equal number of regular and irregular verbs that comprised hard, intermediate and easy verbs. Each trial consisted of the auditory presentation of a single verb stem, following which participants then said out loud the corresponding past tense form as quickly as possible. Participants received ten practice verbs; the practice verbs comprised an equal form as quickly as possible.  Participants received ten practice trials, if necessary, the experimenter intervened with further instructions. Next, the participants experienced the 60 experimental trials. Whereas adult participants completed all 60 trials in one run, the children were given breaks after every 15 trials in order to incorporate a reward recognition scheme: children received a gold star at completion of the practice trials and after every 15 experimental trials which they could stick on a personal ‘certificate of achievement’, resulting in five stars if they completed the whole study. Furthermore, the children were regularly encouraged throughout the study.

**Analysis.** Errors and reaction times were analyzed separately. Errors were scored by listening to the recordings and identifying and classifying incorrect responses. Reaction times were analyzed by manually measuring the duration from the onset of the stimulus to the onset of the participants’ response using the Adobe Audition software (this method was considered more accurate than using a voice key). Additional analyses were performed by measuring reaction times from the middle of the stimulus to the onset of the response to approximate different recognition points for different words; results did not differ systematically from the stimulus onset measures and are not reported here. Only reaction times for correct responses were included in the analysis. Answers with false starts were also excluded. Within each condition, reaction times that fell above or below two standard deviations from the mean reaction time were removed as outliers (the same pattern of results was also observed without removing outliers).

**Results**

Adults produced 94.4% of past tense forms correctly. By contrast, 31.9% of the forms produced by the 10-11 year olds were wrong.

**Response latencies.** Mean response latencies for the 10-11 year olds are shown in Figure 1. A 3x2 ANOVA with conditions easiness (easy, intermediate, hard) and regularity (regular, irregular) revealed a marginal effect of regularity (F(1, 632) = 3.672; p = .056) and no main effect of easiness (F(2, 632) = 1.583 ; p = .206). There was no interaction between easiness and regularity (F(2, 632) = 1.667; p = .190). The marginal effect of regularity was driven by the intermediate verbs for which irregulars took longer to produce than regulars.

These results could be taken as evidence for a dual-mechanism account of verb inflection in which regular and irregular forms are processed differently. From a single mechanism perspective it could be argued that these results might be due to a developing lexicon: as children acquire more words and experience already known words more frequently, the factors determining easiness of an individual word change. This is because the experienced frequencies of a word itself, but also of its friends and enemies changes with experience. Therefore, the overall easiness of a verb changes through life with the changing lexicon. However,
the easiness scores used in this study are largely based on adult norms. According to the DMT the observed pattern of dissociations between regular and irregular verbs should become clearer in adults when both the rule and the associative lexicon have become fully established and the system has separated into two mechanisms. In contrast, a single mechanism view predicts that dissociations between regulars and irregulars should become less pronounced while an easy-hard difference would emerge more clearly.

The mean response latencies for the adult participants are shown in Figure 2. Here, a 3x2 ANOVA revealed no effect for regularity (F(1, 795) = 0.012; p = .912) but a highly significant effect for easiness (F(2, 795) = 10.202; p < .001). There was no interaction between regularity and easiness (F(2, 795) = 0.578; p = .561). Planned comparisons showed that response latencies for hard verbs were significantly slower than for intermediate verbs (t(1, 589) = 2.839; p < .05) and those for intermediate verbs were significantly slower than for easy verbs (t(1, 659) = 2.175; p < .005). These results present strong evidence for a single mechanism processing system in which graded dissociations arise from overall easiness and not from the grammatical status of a verb.

We further investigated which of the factors that had been used in determining overall easiness of a verb correlated with response latencies. Those correlations that were significant at the 0.05 level are displayed in Table 1. Overall, the factors that were significantly correlated with response latencies in children also correlated with adult response times, apart from age of acquisition that was significant for the children but not for adults. Response latencies in adults were more sensitive to the friends and enemies of a verb as well as the frequency of its stem and past tense forms. This result might indicate a better developed lexicon in adults in which different token frequencies have impacted on the representations of individual verbs.

<table>
<thead>
<tr>
<th>Factor</th>
<th>10-11 year olds</th>
<th>Adults</th>
</tr>
</thead>
<tbody>
<tr>
<td>enemy types</td>
<td>n.s.</td>
<td>r = 0.15; n = 463; p = .001</td>
</tr>
<tr>
<td>enemy tokens</td>
<td>n.s.</td>
<td>r = 0.24; n = 463; p &lt; .001</td>
</tr>
<tr>
<td>friend tokens</td>
<td>n.s.</td>
<td>r = 0.11; n = 557; p = .009</td>
</tr>
<tr>
<td>ratio of friend to enemy types</td>
<td>r = -0.12; n = 420; p = 0.015</td>
<td>r = -0.17; n = 572; p &lt; .001</td>
</tr>
<tr>
<td>ratio of friend to enemy tokens</td>
<td>r = -0.11; n = 420; p = 0.031</td>
<td>r = -0.22; n = 572; p &lt; .001</td>
</tr>
<tr>
<td>total number of friend and enemy types</td>
<td>n.s.</td>
<td>r = 0.21; n = 683; p &lt; .001</td>
</tr>
<tr>
<td>total number of friend and enemy tokens</td>
<td>r = 0.11; n = 469; p = 0.017</td>
<td>r = 0.25; n = 668; p &lt; .001</td>
</tr>
<tr>
<td>age of acquisition</td>
<td>r = 0.10; n = 414; p = 0.040</td>
<td>n.s.</td>
</tr>
<tr>
<td>stem final alveolar</td>
<td>r = -0.13; n = 590; p &lt; .001</td>
<td>r = -0.11; n = 787; p = .01</td>
</tr>
<tr>
<td>frequency</td>
<td>n.s.</td>
<td>r = -0.13; n = 787; p &lt; .001</td>
</tr>
<tr>
<td>Past frequency</td>
<td>n.s.</td>
<td>r = -0.17; n = 787; p &lt; .001</td>
</tr>
<tr>
<td>Parental past frequency</td>
<td>n.s.</td>
<td>r = -0.08; n = 607; p = .019</td>
</tr>
</tbody>
</table>

The directions of correlations were as expected: responses were slower for verbs with lower frequency of stem, past tense, and parental input, with more enemy types and
tokens, a lower friends-to-enemy ratio, more total phonological neighbors, a later age of acquisition, and with stem final alveolars.

Despite deserved caution in interpreting correlation data, these results indicate that speed of production of the past tense of a verb is affected by interactions between the friends and enemies of this verb and the verb’s own frequency, suggesting an associative mechanism underlying the production of past tense forms.

**Anti-frequency effect.** Whereas most speeded elicitation studies have reported frequency effects for irregular verbs with more frequent irregular forms being produced more quickly, only some studies have reported an anti-frequency effect for regulars in which low-frequency regular forms are produced faster than higher frequency forms (Prasada et al., 1990; Beck, 1997; Clahsen et al., 2004). Within a dual-mechanism framework this effect has been argued to arise from the storage in the mental lexicon of high-frequency regulars whereas low frequency regulars are produced by the mental rule and thus, faster (Pinker, 1999). In a single-mechanism associative framework, all other factors being equal, one would not expect to find anti-frequency effects because memory-storage improves with increasing frequency, leading to shorter latencies for high-frequency forms compared with low-frequent ones.

![Figure 3](image-url)

**Figure 3:** Adult response latencies for low and high-frequency verbs.

To test for an anti-frequency effect we split the selected verbs into low and high frequency forms based on their frequencies in CELEX. The mean response times for low and high frequency regular and irregular verbs are shown in Figure 3. A 2x2 ANOVA with conditions frequency (low, high) and regularity (regular, irregular) revealed a highly significant frequency effect ($F(1, 783) = 15.730; p < .001$) but no interaction between frequency and regularity ($F(1, 783) = 0.361; p = 0.548$). For both regular and irregular verbs, high-frequency forms were produced faster than lower frequency forms. This result contradicts a dual-mechanism account but supports an account in which all inflections are generated by a single mechanism.

**Errors.** Unsurprisingly, the rate of errors decreased with age. 10-11 year olds made errors for 31.9% of all verbs, and adults for 5.6% of all cases. An additional group of 18 5-6 year old children was tested and displayed an error rate of 57.1%; due to this high rate this group was not analyzed for reaction times. For the 5-6 year olds the most common error for both regulars and irregulars was repetition of the bare stem (35% of regular and 32% of irregular forms) and for irregulars, overregularization (35% of all irregular forms). In 10-11 year olds bare stem errors accounted for 4% of regulars and 5% of irregulars, and 40% of irregulars were overregularized. These children also showed a small irregularization rate of 2% of all regulars. The most common errors in adults were overregularizations (6% of all irregulars) and irregularizations (2% of all regulars). In all age groups, errors were generally higher for harder than for easier verbs (5-6 year olds: easy: 45%, intermediate: 64%, hard: 60%; 10-11 year olds: easy: 18%, intermediate: 39%, hard: 39%; adults: easy: 5%, intermediate: 5%, hard: 7%).

While errors for irregulars were generally higher than for regulars across all age groups, it is important to note that the single mechanism approach does not claim that there is no difference between regular and irregular verbs. The regular case is by far the most common past tense inflection both in types and in tokens. An associative account would therefore predict that irregular forms, especially when similar to existing regulars, would sometimes be prone to overregularization. Irregularizations would be less frequent because clusters of irregulars are much smaller than the large regular group.

**Discussion**

The results presented here indicate that easiness, but not regularity, is the main underlying principle in the processing of past tense inflection in adults. In particular, in a speeded production experiment, adult participants showed no difference between regular and irregular forms, but they showed different response latencies for easy, intermediate and hard verbs. This result supports a single mechanisms view of verb inflection in which differences between verbs are graded and arise from interacting statistical factors that determine the overall processing difficulty for a verb. From this perspective, the processing of a verb depends not only on the characteristics of that verb, but also on its relation to other verbs in the mental lexicon. By contrast, in dual mechanism approaches the way in which a verb is processed is determined by its grammatical class and is largely independent of statistical factors and other verbs. Only in the extended DMT which allows for some regulars to be stored alongside irregulars would a limited effect of statistical factors for regulars be expected. However, as the extended DMT does not provide clear criteria as to which of the regular verbs are stored alongside the irregular forms, it seems only a small step to allow for all verbs to be produced by the same associative mechanism.
The results for 10-11 year olds and adults differ in that the children display a marginal effect of regularity. This can have several reasons. One is, as discussed above, that in their developing lexicon the distributional factors are different from those in adults and that, due to their high type frequency, regular forms take an even more dominant role than in the adult lexicon. Another possible explanation lies in the nature of the task. The children, who were required to generate from the stem of a verb its past tense form, were aware of the rule for regular verbs (one 11 year-old child, when asked “Do you know what the past tense of a verb is?”, replied: “Yeah, it’s when you add –ed”). This may have led to a differential processing of regular and irregular verbs and a high overregularization rate in this age group. However, whereas the DMT predicts that this dissociation becomes stronger as it manifests itself in the adult, a single mechanism view would consider it as temporary explicit knowledge that becomes internalized as the child develops, leading to the disappearance of systematic dissociations between regular and irregular verbs in adults in favor of a graded difference along an easiness continuum. The results presented here support this single mechanism view.

Acknowledgments
This work was supported by ESRC grant RES-061-23-0129 to Gert Westermann. We thank Marcus Priest and Abbie Pringle for helping with data collection and Anna Theakston for providing the parental frequency counts.

References
Prasada, S, Pinker, S, & Snyder, W (1990) Some Evidence that Irregular Forms are Retrieved from Memory but Regular Forms are Rule-generated.