Acceptability Ratings of Regular and Irregular Past-tense Forms: Evidence for a Dual-system Model of Language from Word Frequency and Phonological Neighbourhood Effects

Michael T. Ullman GICCS, Georgetown University, Washington, DC, USA

What are the computational and representational bases of the mental lexicon of words, and of the rules of grammar which productively combine lexical forms into larger words, phrases and sentences? "Dual-system" theories posit that lexical forms with non-compositional (arbitrary) sound-meaning pairings are stored in memory, whereas compositional structures are subserved by a distinct rule-processing system. "Single-system" theories claim that lexicon and grammar are both subserved by a single associative memory. Investigations of English past tense may help to resolve this controversy. On the dual-system view, irregular past-tense forms (e.g. blowblew) are retrieved from memory, whereas regular past-tense forms (e.g. walk-walked) are produced by the application of an -ed-suffixation rule. On the single-system view, both types of past-tense forms are learned and computed in associative memory. To test these competing theories, acceptability ratings were elicited from native English-speaking adults for regular and irregular past-tense forms, and their stems, in sentence contexts. Partialling out stem ratings, ratings of irregular past-tense forms (blew) correlated with their frequencies and with measures of the number of similarsounding irregular verbs (threw, grew), whereas ratings of regular part-tense forms (walked) did not correlate with their frequencies or with measures of the number of similar-sounding regular verbs (stalked, balked). The results suggest that irregular past tenses are retrieved from associative memory, whereas regular past tenses are produced by a suffixation rule.

Requests for reprints should be addressed to Michael T. Ullman, Georgetown Institute for Cognitive and Computational Sciences, Georgetown University, New Research Building, 3970 Reservoir Road NW, Washington DC 20007, USA.

E-mail: michael@giccs.georgetown.edu

I thank Susan Carey, Marie Coppola, Zoubin Ghahramani, Joseph Locascio, Alan Prince and, in particular, Roumyana Izvorski and Steven Pinker, for helpful insights and discussions. This research was carried out while the author was at MIT, Cambridge, MA, and was supported in part by NSF grant BNS 91-09766 and NIH grant HD 18381 to Steven Pinker, and by a Poitras predoctoral fellowship and a McDonnell-Pew postdoctoral fellowship to M.T.U.

INTRODUCTION

Human language is characterised by two capacities. One is the "mental lexicon" of memorised words, a memory store which must contain at least those words with arbitrary sound–meaning pairings (i.e. whose sounds and meanings cannot be derived from each other), such as *walk* or *hand*. The other is the "mental grammar" of rules, which combine forms into predictably structured larger words (e.g. *walked*, *hands*), phrases and sentences, and which productively apply to new words (e.g. *fax–faxed*) and to novel forms (e.g. *blick–blicked*) (Berko, 1958; Chomsky, 1965, 1995; Pinker, 1994).

The two capacities have been explained by two competing theoretical frameworks. "Dual-system" theories posit distinct representational and computational bases for the two capacities. Word sounds and meanings are stored in either a rote or an associative memory. The rules and principles underlying grammatical knowledge and processing are subserved by one or more specialised components that are dependent upon the manipulation of symbolic representations of words and other linguistic structures (Chomsky, 1965, 1995; Fodor, 1983; Frazier, 1987; Pinker, 1994).

"Single-system" theories, on the other hand, posit that lexical items and grammatical rules are represented and processed by a single system. Bates and her colleagues have proposed a model in which lexical and grammatical knowledge is subserved by a large and heterogeneous lexicon (Bates & Goodman, 1997; Bates & MacWhinney, 1989). Similarly, MacDonald, Pearlmutter and Seidenberg (1994, p. 676) hypothesise that "the lexicon is the repository for all types of knowledge associated with words, including their syntactic functions". Modern connectionism (parallel distributed processing) has offered a computational framework for the single-system perspective. On this view, the representation and computation of lexical items and grammatical rules takes place over a large number of interconnected simple processing units. Learning occurs by adjusting weights on connections on the basis of statistical contingencies in the environment. Grammatical rules are nothing but descriptions of behaviour, rather than the mental manipulations posited by dual-system theories (Elman et al., 1996; Rumelhart & McClelland, 1986).

Demonstrations of lexicon/grammar dissociations can strengthen the dual-system view, whereas imperfections in such dissociations can strengthen the single-system view. However, testing for such dissociations has been problematic because tasks probing for lexicon and for grammar usually differ in ways other than their use of the two capacities. For example, it is difficult to match measures of grammatical processing in sentence comprehension with measures of lexical memory (see Bates, Harris, Marchman, Wulfeck, & Kritchevsky, 1995).

One well-studied approach, which largely overcomes this problem, has been the investigation of English past tense. Irregular past-tense transformations are largely unpredictable in sound (e.g. *cling-clung*, *bring-brought*) and do not apply productively (new irregulars rarely enter the language). In contrast, regular past-tense transformations are predictable in sound (verb+-ed) and apply productively to new words and to novel forms (e.g. *fax-faxed*, *blick-blicked*). Thus irregular forms have arbitrary mappings with their stems, just as non-compositional lexical forms have arbitrary mappings with their meanings, whereas the structure of regular past-tense forms can be described by a rule which applies in a productive manner. Crucially, irregulars and regulars are matched in complexity (one word), meaning (past) and syntax (tensed), and can also be matched on syllable structure (e.g. *slept-slipped*), word frequency and other factors.

According to a dual-system view, irregular forms (*blew*) are stored in and retrieved from lexical memory, whereas regular forms (*walked*) are produced in real-time by a distinct symbol-manipulation system which applies an *-ed*-suffixation rule to the verb stem (*walk*). Retrieval of an irregular blocks the rule (*blew* pre-empts *blowed*); when an irregular is not successfully retrieved, the rule may be applied, resulting in "over-regularisation" errors such as *blowed* (Pinker, 1991; Pinker & Prince, 1988).

Two alternative theories have been proposed, motivated in part by the observation that irregular transformations are not completely unpredictable and unproductive, but rather follow subregularities (e.g. springsprang, ring-rang, sing-sang), and can lead to the generation of novel forms (e.g. spling-splang). One theory posits that rules underlie the computation of irregulars (e.g. computing *i–a* in *spring–sprang*, *ring–rang*) as well as regulars (Chomsky & Halle 1968; Halle & Mohanan, 1985; Ling & Marinov, 1993). The second theory is consistent with the single-system view, and claims that regulars and irregulars are learned in and computed over an associative memory. In support of this theory, a number of connectionist models have been developed in which input and output units represent the sounds of verb stems and past-tense forms respectively, and in which the weights of a matrix of input-output connections are adjusted according to how the statistical structure of stem-past pairs influences the behaviour of the network (Daugherty & Seidenberg, 1992; Hare & Elman, 1995; Hare, Elman, & Daugherty, 1995; MacWhinney & Leinbach, 1991; Marchman, 1993; Plunkett & Marchman, 1991, 1993; Rumelhart and McClelland, 1986).

Here, I argue in favour of the dual-system perspective. I present evidence for dissociations between the computation of regular and irregular past-tense forms with respect to their sensitivity to word frequency and to their phonological similarity with other verbs. I argue that this evidence suggests that connectionist models are successful at capturing important aspects of the representation and computation of irregular verbs, but that regular verbs are computed by a distinct rule-processing system.

Evidence suggests that the more often a word is encountered, the better it is remembered (e.g. Rubenstein, Garfield, & Milliken, 1970). According to the dual-system view, irregular past-tense forms (*blew*) are retrieved from memory, so they are expected to be frequency-sensitive, with high-frequency forms being remembered better than low-frequency forms. Regular past-tense forms (*walked*) are rule-produced in real-time, so they should show no such "frequency effects" once access to their stem forms (*walk*), to which the *-ed*-suffixation rule is applied, is controlled for.

In contrast, if the single-system view is correct, and regulars as well as irregulars are memorised in associative memory, then both past-tense types may show word frequency effects (e.g. Daugherty & Seidenberg, 1992). Moreover, many single-system models expect "phonological neighbourhood effects": The memory traces representing the distributed phonological stem–past mappings shared among "neighbouring" irregulars (e.g. sing-sang, ring-rang, spring-sprang) or regulars (e.g. slip-slipped, clip-clipped, trip-tripped) should be strengthened by the learning of any form with these mappings. Thus hearing *sing-sang* should strengthen not only the memory traces unique to *sing-sang*, but also those shared between sing-sang, ring-rang and spring-sprang (e.g. Plunkett & Marchman, 1991, 1993; Rumelhart & McClelland, 1986). Moreover, because there are many regular verbs, and they follow a consistent pattern, it has been claimed that the contribution of neighbouring regulars to a given regular verb's memory traces may largely overwhelm the contribution of the verb's individual past-tense frequency, resulting in weakened past tense frequency effects for regular verbs (Daugherty & Seidenberg, 1992; Seidenberg, 1992).

Although traditional dual-system models assume that irregular past-tense forms are stored in rote memory, more recent models have been influenced by the partial regularity (e.g. spring-sprang, ring-rang, fling-flung) and partial productivity (e.g. spling-splang) of irregulars, and have adopted the single-system connectionist perspective that irregulars are learned in an associative memory (Pinker, 1991). If irregulars are retrieved from associative memory in a manner similar to that hypothesised by single-system models, whereas regulars are rule-products, then phonological neighbourhood effects should be found for irregular but not regular forms. Finally, if irregulars as well as regulars are rule-produced (Halle & Mohanan, 1985; Ling & Marinov, 1993), neither irregular nor regular forms should show either word-frequency or phonological neighbourhood effects.

There have been few studies of word-frequency and phonological neighbourhood effects contrasting regular and irregular past-tense forms. Prasada, Pinker and Snyder (1990) presented adult subjects with a series of verb stems, and asked them to produce the past-tense form of each verb as quickly and accurately as possible. The subjects took significantly more time to produce low-frequency than high-frequency past-tense forms for irregular verbs, holding stem frequency constant. However, they found that production times were not significantly longer for low-frequency than high-frequency regular past-tense forms, holding stem frequency constant. This result is consistent with the hypothesis that irregular but not regular past-tense forms are retrieved from memory. Stemberger and MacWhinney (1988) examined errors on intended past-tense forms in spontaneous speech, and also gave subjects a production task of regular past-tense forms. They reported frequency effects for regulars in the production task. but not in spontaneous speech; however, the frequency effects must be treated with caution, because there was no control for any measure of access to the verb stems, such as stem frequency. Similarly, although Marchman (1997) reported frequency effects for both regular and irregular verbs in a past-tense elicitation task given to children, stem access was not controlled for.

Bybee and Moder (1983) and Prasada and Pinker (1993) investigated phonological neighbourhood effects in novel verbs. Bybee and Moder (1983) presented subjects with novel verbs in obligatory past-tense sentence contexts (Sam likes to ...; Yesterday he ...). The verb stems varied in phonological proximity to the prototypical pattern which the authors defined for the $i-\land$ (e.g. string-strung) group of irregular verbs: sCCV[velar nasal]. They found a continuous effect of similarity as a function of proximity to the prototype: The closer a real or novel verb stem was to the prototype, the more likely it was to be inflected to the \(\lambda\)-form. Thus *spling* was more likely to be inflected as *splung* than *vin* was as *vun*. They took this quasi-productivity as evidence that irregular verbs are organised into family resemblance categories (Rosch, 1978). Prasada and Pinker (1993) replicated Bybee and Moder's (1983) finding that the production of an irregular-sounding past-tense form for a novel verb increases with the similarity of that verb to the prototype of an irregular neighbourhood (spling-splung versus vin-vun). However, Prasada and Pinker found that subjects were not significantly more likely to produce regularly suffixed past-tense forms of familiar-sounding novel verbs like plip than of odd-sounding novel verbs like ploamph—even though the subjects rated the stem *ploamph* to be less acceptable than the stem *plip*. The authors found a similar distinction between acceptability ratings of irregular-sounding (splung) and -ed-suffixed (plipped) past-tense forms of novel verbs: Subjects gave higher ratings to forms like *splung* than *vun*, whereas forms like *plipped* were not given higher ratings than those like *ploamphed*, once the naturalness of their stems was held constant. In her past tense elicitation task given to children, Marchman (1997) reported that errors on irregulars were not predicted by the number of similar irregular neighbours. In contrast, errors on regulars were affected by the number of similar regular neighbours. However, this surprising contrast must be treated with caution, because no statistics were reported for the difference for regulars, and stem access was not controlled for, so any similarity effects of stem forms were not factored out.

There has thus been no systematic psycholinguistic investigation of both frequency and phonological neighbourhood effects of the computation of regular (*walk-walked*) and irregular (*blow-blew*) past-tense forms, holding stem frequency constant. Here, I address this empirical gap, describing a study in which acceptability ratings were elicited from normal adults for regular and irregular past-tense forms, and were analysed for frequency and phonological neighbourhood effects.

METHODS

On a 7-point scale, subjects were asked to give acceptability ratings for past-tense forms of regular and irregular verbs in past-tense sentence contexts, and of their stem forms in present-tense sentence contexts.

Subjects

Each past-tense and stem form was given acceptability ratings by 32 MIT undergraduates, who were paid for their participation in the experiment. All were native speakers of American or Canadian English.

Verbs

Subjects were shown 137 verbs with monosyllabic stems (see Appendix 1 for a full list): 89 were irregular verbs (*blow-blew*) and 48 were regular verbs (*walk-walked*). In addition to these 137 verbs whose ratings are analysed in this paper, the questionnaire contained another 335 verbs not discussed here (see Appendix 2).

The 89 irregular verbs were drawn from Pinker and Prince (1988). Doublet verbs, which have acceptable irregular and regular past-tense forms (*dive-dove/dived*) were excluded. In this study, doublets were defined as those verbs with an irregular past-tense form whose corresponding regularisation's mean acceptability rating was greater than 3.5 out of 7. All regularisations of the non-doublet irregular verbs (*blowed*) had relative frequency counts of 0, according to both the Francis and Kučera (1982) and Associated Press word-frequency counts (see below for

details). The mean (\pm standard deviation) of the In-transformed past-tense frequencies for the 89 irregulars was 3.59 ± 1.55 for the Francis and Kučera frequency counts and 6.63 ± 2.15 for the Associated Press counts.

The 48 "consistent" regulars were drawn from the Francis and Kučera and the Associated Press frequency counts. Their stems were phonologically similar to the stems of other regular verbs, and dissimilar to the stems of irregular verbs, and thus they and their phonological neighbours are consistently regularised (e.g. balk-balked, stalk-stalked). Thus none of their stems rhyme with the stems of irregulars; nor do they have /t/ or /d/ as a final consonant, because many irregular stems end in one of these two phonemes (e.g. wet, bite, ride, bend). The mean In-transformed past-tense frequency for the 48 consistent regulars was 1.77 ± 1.51 for the Francis and Kučera counts and 4.11 ± 2.34 for the Associated Press counts.

Presentation

The position of each past-tense sentence in the questionnaire was randomised, and the order of regulars and irregulars was counterbalanced across subjects. The following is an example sentence:

walk
John and Ralph walked to the store.

1 2 3 4 5 6 7
worst best

Stem naturalness ratings were elicited for all verbs in present-tense sentences in any person other than third person singular. There was no attempt to make the stem sentence of each verb similar in content to that of its past-tense sentence. Each verb stem was presented in a format similar to that in which the past-tense forms were presented. For example:

drop
People often drop things when they have oily hands.

1 2 3 4 5 6 7
worst best

Because ratings were gathered on a large number of verb forms, the full questionnaires were broken down into sub-questionnaires. Each sub-questionnaire contained both the stem and past-tense sentences for a given verb. Each subject completed one or more sub-questionnaires. Thus although each verb form was rated by 32 subjects, not all verbs were rated by the same subject.

Instructions

Before starting the past-tense rating task, subjects were asked to give judgements based on the naturalness of the past-tense form printed in italics in each sentence. The instructions stressed that the experiments were not asking for judgements about the real-world plausibility of the sentences, but rather about the naturalness of the past-tense form in the sentence: "Is the verb in a form that 'sounds' right to you and that you would naturally use in your own speech?" The subjects were asked to use the entire scale between 1 and 7, and it was stressed that "it is important to remember that we are looking for your intuitions and gut feelings, and not what you believe the correct form to be according to what the dictionary says or what your teachers have told you". The instructions preceding the stem rating task were similar to those of the past-tense rating task, requesting the subject to judge the naturalness of the present-tense forms of verbs.

PREDICTORS

If past-tense forms are retrieved from an associative memory representing distributed stem—past mappings, their acceptability ratings may be predicted by their word frequencies as well as by some function of the number (type frequency) and similarity of their phonological neighbours, and of the word frequencies of these neighbours (their token frequencies). In contrast, if past-tense forms are produced by the application of a rule to their stems, their past-tense frequencies and the structure of their phonological neighbours should not affect their acceptability ratings, once a measure of stem access has been held constant. To test these distinctions, measures of past-tense frequency and phonological neighbourhood influences were acquired.

Past-tense Frequency

Two relative frequency counts were used in all word-frequency analyses: (1) frequency counts derived by Francis and Kučera (1982) from 1 million words of text drawn from several sources selected to cover a range of topics; (2) frequency counts extracted by a stochastic part-of-speech analyser from a 44 million word corpus of unedited Associated Press news wires between February and December 1988 (Church, 1988). Hereafter, the two frequency counts are referred to as "FK" and "AP", respectively. Both counts distinguished different parts of speech; for example, *talked* used as a past tense has a separate count from *talked* used as a past participle. All analyses were carried out on the natural logarithm of each raw frequency count, which was first augmented by 1.1 to avoid ln(0).

Neighbourhood Strength

To test the single-system connectionist hypothesis that we learn mappings between distributed phonological representations of stem and past-tense forms in associative memory, and that similar-sounding stem-past pairs (blow-blew, grow-grew, throw-threw) share memory traces representing common mappings (e.g. Plunkett & Marchman, 1991, 1993; Rumelhart & McClelland, 1986), the "neighbourhood strength" function was developed. The function attempts to capture any influence of phonological neighbours on the memory trace strength of a given verb's stem-past mappings.

For each stem-past pair rated by subjects in the experiment (e.g. blow-blew), a neighbourhood strength value was calculated to measure any influence of neighbouring verbs (e.g. throw-threw, grow-grew). For each irregular verb, "irregular neighbourhood strength" was calculated as a function of the phonological structure its neighbouring irregular verbs. For each regular verb (walked), "regular neighbourhood strength" was calculated as a function of the phonological structure of its regular verb neighbours (balked, stalked).

To explain the neighbourhood strength function, let us step through the process of calculating the irregular neighbourhood strength of *blow-blew*. In brief, this value is the sum of the contributions from all irregular verbs with a similar-sounding stem: *grow-grew*, *throw-threw*, *go-went*, etc. Each such neighbouring irregular verb's contribution to the memory traces of *blow-blew* increases with its token frequency and with the phonological similarity of its stem to *blow*. If the two past-tense forms are similar (*blew* and *grew*), then the neighbouring verb is considered to be a "friend", strengthening the memory traces of *blow-blew*. If the the two past-tense forms are dissimilar (*blew* and *went*), then the neighbouring verb is considered to be an "enemy", weakening the memory traces of *blow-blew*.

First, each irregular Neighbour of blow—blew was selected from an online database of 5350 verbs. A Neighbour is defined as having a monosyllabic stem which is phonologically similar to blow, such that $similarity(stem(Neighbour), blow) > \alpha similarity(blow, blow)$. According to the similarity metric used (Tversky, 1977; see below), the similarity between two objects increases with the number of their shared features. Therefore, the similarity of a form with itself increases with the number of phonemes in the form: Longer forms are more self-similar than shorter forms. A reasonable minimum similarity for a Neighbour to affect blow—blew's neighbourhood strength is some percentage of blow's reflexive similarity. The free parameter α was assigned the value 0.25, which was selected by finding a good fit between irregular neighbourhood strengths and the acceptability ratings of a randomly selected subset of irregular

verbs in the experiment. The pattern of results yielded by several other values of α was similar to that reported in this paper.

The greater the phonological similarity between blow-blew and the stem-past pair of a "friendly" Neighbour (e.g. grow-grew), the greater the degree of mutual strengthening of the two pairs: Each shared stem-past feature mapping strengthened in the learning of one pair should also strengthen the representation of the other pair. Conversely, if two verb stems share many features, but those features are mapped to distinct past-tense features (e.g. blow-blew, go-went), then these neighbours are considered to be enemies, and the learning of one pair should weaken the mappings of the other (e.g. Seidenberg, 1992). Given that we do not know which stem features map to which past-tense features, a function was selected which approximated these characteristics. The contribution of the learning of the Neighbour's stem-past mapping to the memory trace mappings of blow-blew was captured as contribution(blow-blew, Neighbour) = similarity(blew, past(Neighbour)) - β similarity(blow, stem(Neighbour)), where β is a free parameter (see below).

Thus, holding stem similarity constant, *contribution* increases positively with increasing similarity between the two past-tense forms, and decreases with decreasing similarity between the two past-tense forms. When similarity between the two past-tense forms decreases to the value of β similarity(blow, stem(Neighbour)), then contribution = 0. When the similarity decreases below this value, contribution becomes negative. Thus this model assumes that the learning of an enemy Neighbour (e.g. go-went) may weaken the representation of blow-blew. The free parameter β was assigned the value of 0.5. Thus we expect the contribution of a Neighbour to be negative, weakening the learned mapping between blow and blew, when the similarity between the Neighbour's past-tense form and blew is less than half of that between the Neighbour's stem and blow. For example, bring-brought may be an enemy of sing-sang.

The *contribution* of the *Neighbour* is weighted by its AP past-tense frequency—the natural logarithm of the AP frequency count, first augmented by 1.1, to avoid $\ln(0)$, and to capture the contribution of verbs with AP frequency counts of 0: *Contribution(blow-blew, Neighbour)* = frequency(past(Neighbour)) (similarity(blew, past(Neighbour)) – β similarity(blow, stem(Neighbour))). Thus as the past-tense frequency of the Neighbour increases, its contribution to blow-blew's memory traces increases, although that contribution can be positive (strengthening its mappings) or negative (weakening its mappings).

The irregular neighbourhood strength of *blow-blew*, and of other irregular verbs, is calculated as the sum of the *contribution* of each irregular *Neighbour*. The same procedure was followed for the calculation of regular neighbourhood strength for regular verbs (e.g. *walk-walked*),

based on monosyllabic neighbouring regular verbs (e.g. *balk–balked*, *stalk–stalked*).

Similarity. The similarity between two past-tense forms or two stem forms was calculated on the basis of the Contrast Model (Tversky, 1977). According to this model, the similarity between two objects or forms a and b is a function of the sum of the features the two forms have in common, minus a function of those features they do not have in common:

$$similarity(a,b) = \theta f(A \cap B) - \alpha f(A - B) - \beta f(B - A)$$

where A and B are the sets of features representing a and b, and θ , α and β are free parameters. Thus two forms can be similar (their similarity value is positive), neutral (their similarity value is 0), or dissimilar (their similarity value is negative).

Similarity values were computed for all pairs of English phonemes. Each phoneme was represented by a set of distinctive features (Chomsky & Halle, 1968). To reflect logarithmic effects widely observed in learning and perception, the natural logarithm was selected as the function f to be applied to the shared and unshared features. The free parameters θ , α and β were assigned values of 1, 4 and 4, respectively, on the basis of informally acquired similarity judgements over a range of phoneme pairs. Thus unshared features have an influence four times that of shared features. The similarity of two phonemes a and b was therefore captured as:

$$similarity(a,b) = \ln(A \cap B) - 4\ln(A - B) - 4\ln(B - A)$$

where A and B are the sets of distinctive features representing phonemes a and b.

These similarity values between all pairs of English phonemes were then used to compute similarities between pairs of stem forms and pairs of pasttense forms. For each pair of verb forms a and b, dynamic programming was used to calculate a maximum similarity between these forms, which were treated as ordered sets of phonemes. All phoneme pairs between the two forms were compared, and a path of maximum similarity (of the examined paths, the one with the largest sum of similarity values) was selected. Where more than one phoneme from one word form had to be "squished" onto a single phoneme on the other word form (e.g. stack, tack), a cost (negative value) was incurred. This cost was set to the maximum similarity between two phonemes—the total number of distinctive features, of which there were 20. Both the phoneme comparison and the "squishing cost" were weighted according to the phonemes' position in their word forms. If both phonemes were in the rhyme portion of the last syllable of their respective words, their weighting was increased

from 1 to 4. In other words, the rhyme portion of the last syllable was weighted four times more than earlier segments of the word. This weighting was assigned on the basis of informally acquired similarity judgements over a range of verb pairs, which underscored the importance of rhyme.

Procedure. All neighbourhood strength values were computed by a program that was written in the C programming language on a Sun workstation. For each verb (e.g. blow-blew) for which irregular or regular neighbourhood strength was calculated, all necessary information (number of syllables, phonetic representations, and frequencies of stem and pasttense forms) of the set of verbs contributing to its neighbourhood strength was extracted from an on-line database of 5350 verbs. This database was constructed from several sources, including the CELEX database, from the Centre for Lexical Information at the University of Nijmegen in the Netherlands, and the FK and AP frequency counts. The pronunciation representations were based on British speech. Although many of these representations did not differ from representations of General American pronunciations, there were some clear distinctions, such as words ending in r (e.g. swear). Phonological similarity values between certain pairs of verb forms may therefore not have been accurate reflections of the phonological similarity of the two verb forms for the subjects tested. However, such inaccuracies should decrease the goodness of fit between neighbourhood strength and subjects' acceptability ratings, suggesting that any associations that are found between the two variables may be even stronger than the analyses would indicate.

RESULTS AND DISCUSSION

Acceptability ratings of irregular past-tense forms (*blew*) correlated positively with their past-tense frequencies, partialling out stem ratings [FK: r(86) = 0.51, P < 0.001; AP: r(86) = 0.56, P < 0.001], or partialling out stem ratings and irregular neighbourhood strength (*threw, grew*) [FK: r(85) = 0.54, P < 0.001; AP: r(85) = 0.58, P < 0.001]. (*P*-values for all correlations are reported as two-tailed.) In contrast, acceptability ratings of regular past-tense forms (*walked*) did not correlate with their frequencies, partialling out stem ratings [FK: r(45) = -0.17, P = 0.241; AP: r(45) = 0.11, P = 0.468], or partialling out stem ratings and regular neighbourhood strength (*balked, stalked*) [FK: r(44) = -0.17, P = 0.260; AP: r(44) = 0.11, P = 0.483]. The acceptability ratings of irregular and regular past-tense forms had similar means (6.69 and 6.70 respectively), standard deviations (0.27 and 0.22) and ranges (5.7–7.0 and 5.8–7.0). Thus differences in

acceptability rating variance between the two groups are unlikely to account for the contrasting correlations.

The contrasting correlations might be explained by the larger sample size of irregular verbs (89 vs 48 regular verbs), or by the larger range of the irregular verbs' past-tense frequencies (ln-transformed AP frequency range of 1.1–13.1 vs 0.1–9.0 for regulars). To control for these factors, 48 irregular verbs whose past-tense forms had ln-transformed AP frequencies less than 9.0 were randomly selected from the larger sample of 89 irregular verbs. Like the larger sample of irregulars, and unlike the regular verbs, the acceptability ratings of the smaller sample of irregular past-tense forms correlated significantly with their past-tense frequencies, partialling out stem ratings [FK: r(45) = 0.38, P = 0.009; AP: r(45) = 0.46, P = 0.001], or partialling out stem ratings and irregular neighbourhood strength [FK: r(44) = 0.40, P = 0.006; AP: r(44) = 0.45, P = 0.002].

Acceptability ratings of irregular past-tense forms (*blew*) correlated with their irregular neighbourhood strengths (*threw*, *grew*), partialling out stem ratings and past-tense frequencies [FK: r(85) = 0.24, P = 0.022; AP: r(85) = 0.24, P = 0.028]. This pattern also held with the same sample of 48 lower-frequency irregular verbs described above [FK: r(44) = 0.42, P = 0.004; AP: r(44) = 0.38, P = 0.009]. In contrast, acceptability ratings of regular past-tense forms (*walked*) did not correlate at all with their regular neighbourhood strengths (*balked*, *stalked*), partialling out stem ratings and past-tense frequencies [FK: r(44) = -0.03, P = 0.859; AP: r(44) = -0.04, P = 0.774].

Connectionist simulations of the learning and computation of the mapping between word spelling and pronunciation, a domain which connectionist modellers have claimed is similar to the learning and computation of the mapping between verb and past-tense form (Plaut, McClelland, Seidenberg, & Patterson, 1996; Seidenberg, 1992), have yielded an interaction between neighbourhood and frequency effects: Higher-frequency items show less of a neighbourhood effect than lower-frequency items (e.g. Seidenberg & McClelland, 1989). This pattern has also been observed empirically in subjects' performance (e.g. Andrews, 1982; Seidenberg, Waters, Barnes, & Tanenhaus, 1984). Thus, if regular as well as irregular past-tense forms are learned in a single-system associative memory, both might be expected to show neighbourhood-strength/word-frequency trade-offs. In contrast, under the dual-system view, this trade-off should be found only for irregular verbs.

The irregular verbs were divided into two groups, one of low past-tense frequency and one of high past-tense frequency, with the median AP past-tense frequency serving as the divider. The regular verbs were likewise median-split into low- and high-frequency groups. Irregular past-tense acceptability ratings (blew) correlated significantly with irregular neigh-

bourhood strength (threw, grew) for the low-frequency verbs [FK: r(41) = 0.33, P = 0.030; AP: r(41) = 0.34, P = 0.026] but not for the high-frequency verbs [FK: r(40) = 0.09, P = 0.574; AP: r(40) = 0.08, P = 0.637], partialling out stem ratings and past-tense frequencies. This pattern also held with the same sample of 48 lower-frequency irregular verbs described above, which were likewise median-split into low- and high-frequency groups: Past-tense acceptability ratings correlated significantly with neighbourhood strength for the low-frequency verbs [FK: r(20) = 0.63, P = 0.002; AP: r(20) = 0.59, P = 0.004] but not for the high-frequency verbs [FK: r(20) = 0.17, P = 0.004] 0.458; AP: r(20) = 0.20, P = 0.373], partialling out stem ratings and pasttense frequencies. In contrast, regular past-tense acceptability ratings (walked) did not correlate at all with their regular neighbourhood strengths (balked, stalked) for either the low-frequency verbs [FK: r(20)] = -0.14, P = 0.532; AP: r(20) = 0, P = 0.957] or the high-frequency verbs [FK: r(20) = 0.06, P = 0.785; AP: r(20) = -0.01, P = 0.966], partialling out stem ratings and past-tense frequencies. The contrast suggests that irregulars, but not regulars, are learned in an associative memory, which can be modelled by connectionist networks that yield frequency/ neighbourhood interactions.

This view is strengthened by the fact that the frequencies of the regular past-tense items (mean 1n-transformed FK: 1.73; AP: 4.10) were lower than those of the full set of 89 irregular past-tense items (FK: 3.59; AP: 6.63), as well as those of the smaller sample of 48 irregular past-tense forms (FK: 3.14; AP: 6.21). Crucially, this contrast also held for the lower frequency half of the 48 regular and 48 irregular verbs, with the former having mean frequencies (FK: 0.59; AP: 2.16) well below those of the latter (FK: 2.04; AP: 4.76), despite the finding that the irregulars and not the regulars showed neighbourhood effects.

It might be claimed that the lack of neighbourhood effects among regulars is consistent with the properties of sigmoid activation functions (e.g. the logistic function) in connectionist models. On this view, because there are a large number of regulars following the same suffixation pattern, the influence of additional regular neighbours on their underlying memory traces should be diminished. Thus differences in neighbourhood strength among regulars might correspond to small differences in acceptability ratings, resulting in weakened neighbourhood strength effects for regular verbs. However, an examination of the data suggested that such an explanation might not be consistent with the finding that there was not even a trend of neighbourhood effects among the regulars. The range of the neighbourhood strengths of the regular items (1717–25124, SD 6303) was considerably larger than the range of the irregular neighbourhood strengths of the full set of 89 irregulars (range –4710–5633, SD 1744) and of the smaller set of 48 irregulars (range –2309–5633, SD 1918). Thus it would

be surprising if no trend at all were found for regular phonological neighbourhood effects. We also examined neighbourhood strength effects in the range of overlap between irregular and regular neighbourhood strengths. Whereas those irregulars with the largest irregular neighbourhood strengths (n = 32, range 25–5633) showed the expected significant positive correlation between neighbourhood strength and past-tense acceptability ratings, partialling out stem ratings and past-tense frequencies [FK: r(28) = 0.48, P = 0.008; AP: r(28) = 0.49, P = 0.006], those regulars with the smallest regular neighbourhood strengths (n = 32, range 1717–20535) did not show any positive correlation at all between neighbourhood strength and acceptability ratings, in the analogous analysis [FK: r(28) = -0.24, P = 0.199; AP: r(28) = -0.26, P = 0.162].

Thus the frequency and neighbourhood strength contrasts shown in this study between irregular and regular verbs are predicted by a dual-system model in which irregular past-tense forms are learned in, and retrieved from, associative memory, whereas consistent regular past-tense forms, whose stems are not phonologically similar to the stems of irregulars, are rule-products. The results cannot be explained by connectionist models that expect any past tense frequency or neighbourhood effects for consistent regular verbs. Moreover, it is unclear whether the lack of even a trend suggesting frequency effects can be explained by connectionist models that expect weakened frequency effects among regulars with large and consistent neighbourhoods (Daugherty & Seidenberg, 1992; Seidenberg, 1992). Finally, the findings also preclude models which claim that irregular as well as regular past-tense forms are rule-produced (Chomsky & Halle, 1968; Halle & Mohanan, 1985, Ling & Marinov, 1993).

Manuscript received October 1997 Revised manuscript received July 1998

REFERENCES

- Andrews, S. (1982). Phonological recoding: Is the regularity effect consistent? *Memory and Cognition*, 10, 565–575.
- Bates, E., & Goodman, J.C. (1997). On the inseparability of grammar and the lexicon: Evidence from acquisition, aphasia and real-time processing. *Language and Cognitive Processes*, 12, 507–584.
- Bates, E., Harris, C., Marchman, V., Wulfeck, B., & Kritchevsky, M. (1995). Production of complex syntax in normal aging and Alzheimer's disease. *Language and Cognitive Processes*, 10, 487–539.
- Bates, E., & MacWhinney, B. (1989). Functionalism and the competition model. In B. MacWhinney & E. Bates (Eds), *The crosslinguistic study of sentence processing* (pp. 3–76). New York: 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.
- Comsky, N. (1965). Aspects of the theory of syntax. Cambridge, MA: MIT Press.
- Chomsky, N. (1995). The minimalist program. Cambridge, MA: MIT Press.
- Chomsky, N., & Halle, M. (1968). *The sound pattern of English.* New York: Harper & Row.
- Church, K. (1988). A stochastic parts program and noun phrase parser for unrestricted text. Paper presented at the Second Conference on Applied Natural Language Processing, Austin, TX.
- Daugherty, K., & Seidenberg, M. (1992). Rules or connections? The past tense revisited. In Proceedings of the Fourteenth Annual Conference of the Cognitive Science Society (pp. 259–264). Hillsdale, NJ: Lawrence Erlbaum Associates Inc.
- Elman, J., Bates, E., Johnson, M., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1996).
 Rethinking innateness: A connectionist perspective on development. Cambridge, MA: MIT Press.
- Fodor, J. (1983). Modularity of mind. Cambridge, MA: MIT Press.
- Francis, N., & Kučera, H. (1982). Frequency analysis of English usage: Lexicon and grammar. Boston, MA: Houghton Mifflin.
- Frazier, L. (1987). Sentence processing: A tutorial review. In M. Coltheart (Ed.), *Attention and performance XII* (pp. 559–586). Hove, UK: Lawrence Erlbaum Associates Ltd.
- Halle, M. & Mohanan, K.P. (1985). Segmental phonology of modern English. *Linguistic Inquiry*, 16, 57–116.
- Hare, M., & Elman, J. (1995). Learning and morphological change. Cognition, 56, 61-98.
- Hare, M., Elman, J.L., & Daugherty, K.G. (1995). Default generalization in connectionist networks. *Language and Cognitive Processes*, 10, 601–630.
- 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.
- Ling, C., & Marinov, M. (1993). Answering the connectionist challenge: A symbolic model of the past tense acquisitions. *Cognition*, 49, 235–290.
- MacDonald, M.C., Pearlmutter, N.L., & Seidenberg, M.S. (1994). Lexical nature of syntactic ambiguity resolution. *Psychological Review*, *101*, 676–703.
- MacWhinney, B., & Leinbach, J. (1991). Implementations are not conceptualizations: Revising the verb learning model. *Cognition*, 40, 121–157.
- Marchman, V.A. (1993). Constraints on plasticity in a connectionist model of the English past tense. *Journal of Cognitive Neuroscience*, 5, 215–234.
- Marchman, V.A. (1997). Children's productivity in the English past tense: The role of frequency, phonology, and neighbourhood structure. *Cognitive Science*, 21, 283–304.
- Pinker, S. (1991). Rules of language. Science, 253, 530-535.
- Pinker, S. (1994). The language instinct. New York: William Morrow.
- Pinker, S., & Prince, A. (1988). On language and connectionism: Analysis of a parallel distributed processing model of language acquisition. *Cognition*, 28, 73–193.
- Plaut, D.C., McClelland, J.L., Seidenberg, M.S., & Patterson, K. (1996). Understanding normal and impaired word reading. Computational principles in quasi-regular domains. *Psychological Review*, *103*, 56–115.
- Plunkett, K., & Marchman, V. (1991). U-shaped learning and frequency effects in a multi-layered perceptron: Implications for child language acquisition. *Cognition*, 38, 43–102.
- Plunkett, K., & Marchman, V. (1993). From rote learning to system building: Acquiring verb morphology in children and connectionist nets. *Cognition*, 48, 21–69.
- Prasada, S., & Pinker, S. (1993). Generalisation of regular and irregular morphological patterns. *Language and Cognitive Processes*, 8, 1–56.

- Prasada, S., Pinker, S., & Snyder, W. (1990). Some evidence that irregular forms are retrieved from memory but regular forms are rule-generated. Paper presented at the 31st Annual Meeting of the Pschyonomics Society, New Orleans, LA, 16–18 November.
- Rosch, E. (1978). Principles of categorization. In E. Rosch & B.B. Lloyd (Eds), *Cognition and categorization* (pp. 27–48). Hillsdale, NJ: Lawrence Erlbaum Associates Inc.
- Rubenstein, H., Garfield, L., & Milliken, J.A. (1970). Homographic entries in the internal lexicon. *Journal of Verbal Learning and Verbal Behavior*, 9, 487–492.
- Rumelhart, D.E., & McClelland, J.L. (1986). On learning the past tenses of English verbs. In J.L.McClelland, D.E. Rumelhart, & the PDP Research Group (Eds), *Parallel distributed processing: Explorations in the microstructures of cognition*, Vol. 2 (pp. 216–271). Cambridge, MA: Bradford/MIT Press.
- Seidenberg, M. (1992). Connectionism without tears. In S. Davis (Ed.), *Connectionism: Theory and Practice* (pp. 84–137). New York: Oxford University Press.
- Seidenberg, M.S., & McClelland, J.L. (1989). A distributed, developmental model of word recognition and naming. Psychological Review, 96, 523–568.
- Seidenberg, M.S., Waters, G.S., Barnes, M.A., & Tanenhaus, M.K. (1984). When does irregular spelling or pronunciation influence word recognition? *Journal of Verbal Learning and Verbal Behavior*, 23, 383–404.
- Stemberger, J., & MacWhinney, B. (1988). Are inflected forms stored in the lexicon? In M. Hammond & M. Noonan (Eds), *Theoretical morphology: Approaches in modern linguistics* (pp. 101–116). New York: Academic Press.
- Stromswold, K.J. (1990). Learnability and the acquisition of auxiliaries. *Dissertation Abstracts International*, 52, 2535.
- Tversky, A. (1977). Features of similarity. Psychological Review, 84, 327–352.

APPENDIX 1: STIMULI

The following two tables show the stem and past-tense forms of the regular and irregular verbs whose elicited acceptability ratings are reported in this paper. Each Francis and Kučera (1982) and Associated Press relative frequency count is natural logarithm transformed after being augmented by 1.

Irregular verbs

Stem		Irregular Past Tense				
Form	Rating	Form	Rating	FK Freq	AP Freq	
bleed	6.90	bled	6.87	1.09	2.70	
breed	6.90	bred	6.68	0.00	2.56	
feed	6.90	fed	6.78	2.19	4.46	
lead	6.78	led	6.75	4.41	8.37	
read	6.93	read	6.81	3.61	6.94	
meet	6.84	met	6.90	4.39	8.39	
hide	6.93	hid	6.65	1.94	5.34	
slide	6.81	slid	6.62	3.21	5.22	
bite	6.71	bit	6.59	2.07	4.21	
shoot	6.96	shot	7.00	2.94	7.61	
bend	6.93	bent	6.75	2.70	3.98	
send	6.87	sent	6.78	4.24	8.13	
spend	6.71	spent	6.59	3.71	8.05	
lend	6.56	lent	6.53	1.38	4.80	
build	6.90	built	6.90	3.09	6.61	
lose	7.00	lost	6.93	3.91	8.27	
deal	6.81	dealt	6.65	2.19	5.53	
feel	6.93	felt	6.84	5.71	7.93	
mean	6.65	meant	6.75	4.26	6.73	
keep	6.93	kept	6.93	4.75	7.65	
sleep	6.90	slept	6.96	2.94	5.21	
sweep	6.65	swept	6.81	2.99	6.19	
leave	6.75	left	6.65	5.06	8.86	
buy	6.75	bought	6.78	3.49	7.56	
bring	6.81	brought	6.81	4.89	8.06	
catch	6.62	caught	6.65	4.00	6.64	
fight	7.00	fought	6.78	3.17	6.83	
seek	6.87	sought	6.87	3.58	7.76	
teach	6.75	taught	6.90	2.99	6.04	
think	6.75	thought	6.81	5.83	8.49	
flee	6.78	fled	6.84	3.13	7.40	
say	6.90	said	6.78	7.46	13.07	
hear	6.96	heard	6.78	4.86	7.67	
make	6.62	made	6.90	6.14	9.36	
sell	6.81	sold	6.87	3.04	7.67	
tell	6.71	told	6.87	5.65	10.18	
freeze	6.93	froze	6.78	0.69	4.54	
speak	6.96	spoke	6.84	4.46	8.60	
steal	6.81	stole	6.62	2.39	6.13	
get	6.50	got	6.50	5.82	9.31	
swear	6.71	swore	6.81	2.70	4.30	
tear	6.68	tore	6.93	2.77	5.74	

Irregular verbs

Stem		Irregular Past Tense				
Form	Rating	Form	Rating	FK Freq	AP Freq	
wear	7.00	wore	6.90	4.18	6.68	
bear	6.46	bore	6.00	2.70	5.48	
break	6.65	broke	6.75	4.20	8.24	
choose	6.84	chose	6.81	3.63	7.03	
ring	6.81	rang	6.59	3.09	5.02	
sing	6.75	sang	6.53	3.36	6.61	
spring	6.75	sprang	6.06	2.63	4.24	
shrink	6.81	shrank	5.93	0.69	5.23	
sink	6.90	sank	6.15	2.94	6.53	
swim	6.90	swam	6.62	1.94	5.01	
run	6.96	ran	6.87	4.90	8.01	
sit	7.00	sat	6.84	4.94	7.09	
cling	6.78	clung	6.68	2.63	4.00	
fling	6.62	flung	6.59	2.30	3.17	
sling	6.71	slung	5.90	0.69	1.60	
sting	6.75	stung	6.37	0.69	2.30	
string	6.68	strung	6.00	0.69	1.94	
swing	6.87	swung	6.40	3.78	4.39	
wring	6.34	wrung	5.71	0.00	1.09	
hang	6.87	hung	6.43	3.98	5.64	
stick	6.90	stuck	6.53	2.63	5.22	
strike	6.90	struck	6.53	3.71	7.23	
dig	6.59	dug	6.50	2.07	4.69	
win	6.93	won	6.84	3.82	8.80	
blow	6.87	blew	6.87	2.56	6.87	
grow	6.81	grew	6.56	4.18	7.53	
know	6.93	knew	6.87	5.97	8.15	
	6.96		6.78	3.85	7.12	
throw		threw drew	6.87		7.12	
draw	7.00			4.15		
fly	6.96 6.71	flew	6.78 6.68	3.33	7.52 8.46	
see		saw		5.82		
take	6.90	took	6.84	6.05	9.44	
shake	6.84	shook	6.43	4.06	6.32	
bind	6.62	bound	6.31	1.79	3.95	
find	6.65	found	6.96	5.59	9.06	
rise	6.75	rose	6.81	4.11	9.24	
write	6.93	wrote	6.81	5.19	8.50	
ride	6.84	rode	6.59	3.71	5.78	
stride	6.59	strode	6.50	2.39	3.93	
drive	6.84	drove	6.81	4.07	7.22	
fall	6.59	fell	6.87	4.47	9.17	
hold	6.87	held	6.93	4.83	8.41	
come	6.62	came	6.93	6.42	9.52	
eat	6.93	ate	6.96	2.83	5.78	
give	6.90	gave	6.93	5.65	9.00	
stand	7.00	stood	6.84	5.29	7.59	
go	6.96	went	6.96	6.23	9.20	

Regular verbs

Stem		Regular Past Tense				
Form	Rating	Form	Rating	FK Freq	AP Freq	
chop	6.71	chopped	6.87	0.69	3.13	
drop	6.87	dropped	6.90	4.34	8.13	
flop	6.90	flopped	6.59	1.94	3.29	
hop	6.87	hopped	6.84	1.79	3.68	
lop	6.84	lopped	6.46	0.69	0.69	
sop	6.68	sopped	6.43	0.00	0.00	
stop	6.93	stopped	6.59	4.64	7.62	
cross	6.84	crossed	6.81	3.29	6.20	
toss	6.93	tossed	6.87	3.13	5.37	
clap	6.90	clapped	6.87	1.60	3.91	
flap	6.87	flapped	6.84	1.60	2.07	
lap	6.68	lapped	6.75	0.69	1.38	
slap	6.87	slapped	6.81	1.94	4.57	
jar	6.28	jarred	6.81	0.69	2.30	
mar	6.09	marred	6.71	0.00	3.29	
spar	6.34	sparred	6.50	0.00	2.56	
balk	6.65	balked	6.78	1.09	4.67	
stalk	6.43	stalked	6.68	1.94	3.29	
talk	6.93	talked	6.78	3.73	7.17	
walk	7.00	walked	6.68	4.96	7.35	
aim	7.00	aimed	6.87	2.39	5.57	
blame	7.00	blamed	6.75	1.79	7.36	
claim	6.87	claimed	6.96	3.25	8.26	
maim	6.56	maimed	6.75	0.00	1.60	
cram	6.96	crammed	6.90	0.00	1.94	
ram	6.56	rammed	6.78	1.38	4.71	
slam	6.78	slammed	6.65	2.63	5.64	
ask	6.75	asked	6.75	5.70	9.04	
bask	6.81	basked	6.73	0.69	3.33	
bore	6.37					
		bored	6.56	1.38	1.94	
gore	6.06	gored	6.00	0.00	1.60	
pore	5.56	pored	5.84	0.69	1.94	
soar	6.93	soared	6.71	1.38	6.08	
scour	6.56	scoured	6.71	0.00	3.33	
prowl	6.90	prowled	6.50	0.69	1.94	
scowl	6.56	scowled	6.65	1.60	0.00	
chase	6.75	chased	6.87	0.00	5.59	
race	6.90	raced	6.59	2.48	5.25	
trace	7.00	traced	6.93	1.09	4.51	
cook	6.37	cooked	6.62	1.09	3.63	
look	6.81	looked	6.37	5.79	7.63	
dash	6.75	dashed	6.68	1.60	4.07	
gnash	6.43	gnashed	6.37	0.00	0.00	
slash	7.00	slashed	6.81	1.94	4.65	
crush	6.96	crushed	6.96	1.09	5.30	
flush	6.84	flushed	6.75	0.69	1.79	
gush	6.93	gushed	6.81	1.79	2.94	
rush	6.96	rushed	6.84	3.04	6.20	

APPENDIX 2: EXCLUDED STIMULI

The full questionnaires contained not only the 89 irregular and 48 regular past-tense forms whose ratings are analysed in this paper, but the past-tense forms of an additional 335 verbs as well. These verbs comprised several categories which were omitted for a variety of reasons:

- 1. There were 20 monosyllabic "doublet" verbs, like *dive-dove/dived*, which take both an irregular and a regular past-tense form.
- 2. There were 111 monosyllabic "inconsistent regular" verbs, whose stems are phonologically similar to the stems of irregular verbs (e.g. <code>glide-glided</code>; cf. <code>ride-rode</code>, <code>hide-hid</code>). Unlike "consistent regulars", whose stems are phonologically dissimilar to the stems of regular verbs, inconsistent regular verbs are predicted by a dual-system view to be memorised; if they were not, people would utter forms like <code>glid</code> or <code>glode</code>, which does not appear to occur. They will be discussed in a future paper.
- 3. There were 120 prefixed or other multisyllabic verbs (these 120 verbs do not include the three past-tense forms *forbid*, *forbad*, *forbade*, which are discussed below): 61 irregular or doublets verbs (e.g. *forgo-forwent/forgoed*, *interweave-interwove/interweaved*), 18 consistent regulars (e.g. *devour-devoured*, *overlook-overlooked*) and 41 inconsistent regulars (e.g. *accost-accost/accosted*, *decide-decid/decided*). These verbs were excluded from analysis because this study focuses on monosyllabic verbs.
- 4. There were four past-tense forms which can take the role of auxiliary or modal (*be-was*, *be-were*, *have-had*, *do-did*). These were treated separately because they may have a distinct linguistic and psychological status (e.g. Stromswold, 1990).
- 5. There were 11 irregular past-tense forms belonging to five verbs with more than one irregular past, as determined by the author's dialect or by FK and AP frequency counts greater than 0 for more than one of the irregular past-tense forms: spit-spit/spat, bid-bid/bade, forbid-forbid/forbade/forbade, drink-drank/drunk and stink-stank/stunk.
- 6. There were 18 verbs which the author judged to be possible de-nominals, de-adjectivals or of onomatopoeic origin: 8 consistent regulars (e.g. *place-placed*) and 10 inconsistent regulars (e.g. *trust-trust/trusted*). These forms were excluded because their computation may differ from that of forms not derived from another grammatical category (Kim, Pinker, Prince, & Prasada, 1991).
- 7. There were four verbs whose stems or past-tense forms seemed to be easily confoundable with another word (e.g. with the consistent regular *found–founded*, the stem can be confused with the irregular past tense *found*).
- 8. There were 18 no-change irregulars: 13 irregulars (e.g. *hit-hit*) and 5 doublets (e.g. *wet-wet*). These were eliminated because their stem and past-tense forms can be confounded, and because the FK and AP frequency counts may not have reliably distinguished between present and past-tense uses of these verbs.
- 9. There were five verbs whose irregular past-tense forms are marginally acceptable in American English: *learn-learnt/learned*, *spell-spelt/spelled*, *smell-smelt/smelled*, *spoil-spoilt/spoiled* and *cleave-cleft/cleaved*.
- 10. There were seven verbs which were not originally classed as doublets, but whose low irregular ratings and high regular ratings suggested doublethood (e.g. *smite-smote/smited*, *weave-wove/weaved*). They were classified neither as doublet nor as non-doublet irregulars.
- 11. There were 11 verbs whose stems have a final /t/ or /d/, and are thus similar to the stems of many irregulars, but were originally categorised as consistent regulars (e.g. *chat-chatted*, *boast-boasted*, *shout-shouted*). They were analysed neither as inconsistent regulars nor as consistent regulars.
- 12. There were six verbs with accidental errors in their presentation or apparent errors in their frequency counts.