LANGUAGE AND COGNITIVE PROCESSES, 1993, 8 (1) 1–56

P 37 iL

Generalisation of Regular and Irregular Morphological Patterns

Sandeep Prasada and Steven Pinker

Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA

Both regular inflectional patterns (*walk–walked*) and irregular ones (*swing–swung*) can be applied productively to novel words (e.g. *wug–wugged*; *spling–splung*). Theories of generative phonology attribute both generalisations to rules; connectionist theories attribute both to analogies in a pattern associator network; hybrid theories attribute regular (fully predictable default) generalisations to a rule and irregular generalisations to a rote memory with pattern-associator properties. In three experiments and three simulations, we observe the process of generalising morphological patterns in humans and two-layer connectionist networks. Replicating Bybee and Moder (1983), we find that people's willingness to generalise from existing irregular verbs to novel ones depends on the global similarity between them (e.g. *spling* is readily inflectable as *splung*, but *nist* is not inflectable as *nust*). In contrast, generalisability of the regular suffix does not appear to depend on similarity to existing regular verbs: Regularly suffixed versions of both common-sounding *plip* and odd-sounding *ploamph* were reliably produced and highly rated, and the odd-sounding verbs were not rated as having worse past-tense forms, relative to the naturalness of their stems, than common-sounding ones. In contrast, Rumelhart and McClelland's connectionist past-tense model was found to vary strongly in its tendency to supply both irregular and regular inflections to these novel items as a function of their similarity to forms it was trained on, and for the dissimilar forms, successful regular

---

Requests for reprints should be addressed to Steven Pinker, Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, E10–016, Cambridge, Massachusetts 02139, USA. Prasada is now at the Institute for Research in Cognitive Science at the University of Pennsylvania, Philadelphia, PA, USA.

This research was supported by NIH Grant HD 18381 and NSF Grant BNS 91–09766 to the second author and by the McDonell-Pew Center for Cognitive Neuroscience at MIT. The first author was supported by NIH Training Grant T32-MH 18823 and by a fellowship funded by a grant from the Fairchild Foundation to MIT. We would like to thank Alan Prince, John Kim, Gary Marcus, Micheal Ullman, Robbie Jacobs, John Houde, Philip Sabes and three anonymous reviewers for their comments on an earlier draft of the manuscript. Also, thanks to Clayton MacMillan and Paul Smolensky for supplying us with the computer simulation.

inflection rarely occurred. We suggest that rule-only theories have trouble explaining patterns of irregular generalisations, whereas single-network theories have trouble explaining regular ones; the computational demands of the two kinds of verbs are different, so a modular system is optimal for handling both simultaneously. Evidence from linguistics and psycholinguistics independently calls for such a hybrid, where irregular pairs are stored in a memory system that superimposes phonological forms, fostering generalisation by analogy, and regulars are generated by a default suffix concatenation process capable of operating on any verb, regardless of its sound.

INTRODUCTION

Any useful language mechanism must provide the means for productivity —the ability to encode and decode an unlimited number of novel meanings in words and sentences. Traditionally, this ability has been explained by attributing to language users knowledge of grammatical rules that concatenate symbols to form-derived linguistic representations. A simple example is the English regular past-tense rule, which adds the suffix *-ed* to the end of a verb to indicate that the event referred to by the verb took place before the speech act (e.g. *walk–walked*). The suffixation rule, because it can apply to any verb whatsoever, affords unlimited productivity: Given a new verb, speakers can create a past-tense form for it, or understand a past-tense form used by others, without ever having heard it used in the past tense. Such an ability is commonly seen in neologisms (e.g. *faxed*, *xeroxed*) and has been experimentally demonstrated in children as young as 4, who, given a novel verb *to rick*, easily produce *ricked* to refer to past instances (Berko, 1958).

Many languages, though, have forms that are unpredictable and hence cannot be generated by a general rule. For example, English contains about 180 verbs that have "irregular" past-tense forms, such as *hit–hit*, *sing–sang*, *catch–caught*, *go–went*. Traditionally, it has been thought that these idiosyncratic forms are learned by rote and stored as a list in lexical memory. This is consistent with the fact that irregulars tend to be high in frequency, hence more easily memorised, and that the lower-frequency forms among them tend to be erred on more often by children and adults (Bybee & Slobin, 1982; Marcus et al., 1992; Slobin, 1971), and to become regular over historical time-spans (Bybee, 1985). Retrieval of an irregular form blocks application of the regular rule; if no form is retrieved, the rule always applies.

Upon more detailed study, the equation of regular inflection with productive rules, and irregular verbs with unproductive rote listing, turns out to be incorrect. It has long been noted that irregular alternations such as vowel changes are usually found not in single irregular verbs but in families of similar verbs, such as *sing*, *ring*, *spring*, *drink*, *shrink*, *sink*, *stink*, *begin*, *swim*; *keep*, *creep*, *deal*, *feel*, *kneel*, *mean*, *dream*, *grow*, *blow*, *throw*, *know*, *draw*, *fly*; *hit*, *fit*, *split*, *set*, *let*, *put*, and so on. Rote memory, if conceived of as a list of slots, seems unable to explain why this large degree of redundancy is maintained (Pinker & Prince, 1988). Moreover, irregular pairs are not a completely unproductive fixed list; their patterns can to some degree be extended to new forms on the basis of similarity to existing forms. Children occasionally make errors such as *bring–brang* and *bite–bote* (see Bybee & Slobin, 1982; Marcus et al., 1992; Pinker & Prince, 1988). Irregulars can be added to the language diachronically by analogy with existing forms (e.g. *knelt* and *quit*), a process that is especially obvious when dialects are compared (e.g. *bring–brung*, *drag–drug*, *climb–clumb*, *heat–het*; see Mencken, 1936). Finally, adult speakers of the standard dialect frequently extend irregular patterns to nonce stems in experimental tests (e.g. *spling–splung*; Bybee & Moder, 1983; Kim, Pinker, Prince, & Prasada, 1991b).

In some theories of generative grammar (e.g. Chomsky & Halle, 1968; Halle & Mohanan, 1985), irregular patterns are captured in rules (e.g. "change *i* to *a*") just like the regular pattern. Such theories, however, do not adequately capture the patterns of generalisation seen in the irregular system. If an irregular rule is simply attached to the list of words it applies to, the similarity among the words is unexplained. But if a rule is attached to some *pattern* shared by the words (e.g. "change *i* to *a* following a consonant cluster and preceding *ng*"), it fails because the similarity to be accounted for is one of *family resemblance* rather than necessary or sufficient conditions. The rule, while successfully embracing *spring, shrink, drink*, would incorrectly include *bring–brought* and *fling–flung* and would fail to include *begin–began* and *swim–swam* (Bybee & Slobin, 1982).

Moreover, people appear to treat irregular patterns as a kind of graded analogy. Bybee and Moder (1983) asked subjects to produce the past-tense forms of nonsense verbs varying in phonological proximity to the prototypical pattern for *i*-ʌ vowel change verbs. They found a continuous effect of similarity. For example, 44% of the subjects converted *spling* to *splung* (cf. *string*, *sling*, *fling*, *cling*, etc.), 24% converted *shink* to *shunk*, and only 7% converted *sid* to *sud*. Bybee and Moder suggest that people can treat irregular patterns somewhat productively, not with a rule, but with a prototype schema for the related irregular verbs stored in memory, for example, *s*CCV[velar nasal] for the *string* family. The better a verb matches the prototype, the more likely a speaker will productively extend an irregular pattern to it.

Given that irregular patterns are partly productive, but perhaps not rule-generated, a natural question arises: Are rules necessary at all, even for regular verbs? Bybee and Moder suspected they might not be. Like the irregulars, regular verbs might be generalised according to their graded similarity to prototypes, rather than across-the-board application of a rule; Bybee and Moder noted that there were no data relevant to this issue. Their conjecture has been given new plausibility by the connectionist (parallel distributed processing) model of Rumelhart and McClelland (1986). The model has a single associative network consisting of a pool of input units, a pool of output units and a matrix of modifiable weighted links between every input unit and output unit. The stem is represented by turning on a subset of input nodes, each one corresponding to a sequence of three phonological features in the stem. This sends a signal across each of the links to the output nodes, which represent the phonological feature sequences contained in the past-tense form. Each output node sums its incoming signals and turns on if the sum exceeds a threshold; an output module reconstructs the output form by settling on the phoneme sequence that is most compatible with the set of active output nodes. During the learning phase, the past-tense form computed by the network is juxtaposed with a correct version supplied by a "teacher", and the strengths of the links and thresholds are adjusted so as to reduce the difference. There are no units corresponding to words; patterns of alternation between different stem–past pairs are superimposed. This allows the model to generalise automatically to new forms to the extent that their sounds overlap with old ones. The generalisation phenomenon is qualitatively the same for regulars and irregulars: *stopped* is produced because input *op* units were linked to output *opped* units by previous verbs; *clung* is produced because *ing* was linked to *ung*. As a result, the net could model people's analogising of irregular patterns to new forms, while suggesting that such a mechanism is sufficient for regular, seemingly rule-generated productivity as well. Indeed, Rumelhart and McClelland and others have suggested that this model offers a radical alternative to theories of language that explain linguistic abilities in terms of psychologically implemented rules.

In sum, by upsetting the traditional equations of irregularity and lexical listing, and regularity and general rule, the work of Bybee and of Rumelhart and McClelland has left the proper explanation of the regular–irregular distinction, and the psychological status of rules, unclear. Subtler differences in the extent and kind of productivity must be considered.

One of these differences is the manner in which analogy-based and rule-based models generalise knowledge from known items to novel items. The model of Rumelhart and McClelland did not generalise the regular mapping across the board to new items: It failed to supply the correct past-tense form for 33% of the new regular verbs it was tested on. Rumelhart and McClelland noted that these generalisation failures could serve as an empirical test of their model, and that current psychological data were not inconsistent with it and might even support it:

> It is true that the model does not act as a perfect rule-applying machine with novel past-tense forms. However, it must be noted that people – or at least children, even in early grade-school years – are not perfect rule-applying machines either. For example, in Berko's classic study (1958), her kindergarten and first-grade subjects did often produce the correct past forms of novel verbs like *spow*, *mott*, and *rick*, but they did not do so invariably. In fact, the rate of regular past-tense forms given to Berko's novel verbs was only 51 percent. Thus, we see little reason to believe that our model's "deficiencies" are significantly greater than those of native speakers of comparable experience. (pp. 265–266)

Thus it is important to compare the model's generalisation failures with those of humans. Among the regular verbs that the model did not suffix properly, there are three interesting types. One reveals an in-principle limitation of the model, discussed in detail by Pinker and Prince (1988) and Kim et al. (1991b). It hinges on the fact that the model represents only the phonological properties of verb stems, not their morphological structure, and hence systematically errs on irregular-sounding verbs derived from nouns. We will turn to this problem in the General Discussion. The other two generalisation failures motivate the current experiments.

Competition Between Irregular Patterns and the Regular Process

Some regular test verbs resembled irregular stems or pasts, and elicited responses appropriate to the corresponding irregular. For example, the model failed to inflect several verbs ending in *t* or *d* (e.g. *kid–kid*), presumably because of the pattern displayed in many English irregulars that undergo no change in the past tense and that end with *t* or *d* (e.g. *hit–hit*, *rid–rid*). Occasionally, an inappropriate vowel change would also be extended to a novel regular resembling an irregular, as in *sip–sepped* and *slip–slepped* (cf. *sleep–slept*). This effect probably has some psychological reality. Many of Berko's stimuli had the unplanned property of resembling irregular verbs or their past forms (compare, e.g. *spow–throw*, *rick–stick*, *mott–got*, *bodd–bought*), presumably competing with the regular response (see Brown, 1973, for an extensive discussion). Rumelhart and McClelland (1986), Stemberger (1990), Ullman and Pinker (1990; 1991) and Marcus et al. (1992) note several other effects whereby similarity to an irregular pattern inhibits regular suffixation in children and adults. This clearly shows the effect of irregular patterns being stored in some kind of associative memory, which causes new items partly matching the stored patterns to elicit the appropriate irregular form, blocking (or partly blocking) the regular process.

Similar effects can be seen in the model's response to new *irregular* verbs. It is not clear what should count as an "error" for such verbs; *we* know that they are irregular in English, but the model (or the child), lacking access to the "correct" past-tense form, has no way of knowing that. The model produced regularised versions for 11 of the 14 new irregulars but in addition produced no-change forms for *bid, wind* and *grind* (all *d*-final), and the correct English irregular forms as second choices for *weep* and *cling*. Once again, the model shows an ability to generalise by similarity, an ability that seems to have some counterpart in the psychology of language.

Phonologically Unfamiliar Roots

Analogy-based models generalise an input–output mapping to a new item only to the extent that the new item shares features with (i.e. is similar to) previously learned items undergoing that mapping. Since the only continuous dimension of similarity relevant to the past-tense mapping is the phonology of the stem (see Kim et al., 1991b; Pinker & Prince, 1988; Pinker, 1991; and the General Discussion), Rumelhart and McClelland used phonological features to define their distributed representation and thus benefited from appropriate kinds of generalisation due to similarity. For the same reason, however, the model performed poorly in generalising to new (non-trained) regular verbs that differed from any of the regular or irregular verbs in the training set. For the verbs *jump, pump, soak, warm, trail* and *glare*, which do not resemble irregulars, the model produced no suprathreshold output at all, and for certain words related to these, it yielded odd blends such as *mail–membled* and *tour–toureder*. In comparison, if humans possess a symbolic rule that appends the *-ed* suffix to any word classified as an example of the abstract category "verb stem" (unless based on an irregular root), they should be capable of producing a regular past tense for any new verb, regardless of whether it sounds similar to previously encountered regular verbs. Thus similarity-driven and rule-based models would appear to differ in their predictions about humans' ability to inflect verbs with very novel sound patterns.

Pinker and Prince (1988) suggested that people's ability to generate regular past forms appears to extend to all sound patterns, no matter how unusual. For example, any adult can inflect unusual sounds like *stint, anastomose, incommode* and *prescind*. But this observation is informal and in need of further investigation. Even rare words may share enough features with existing regulars to foster generalisation (e.g. *print* for *stint*). Moreover, very few features are necessary if they have uniformly been associated with the regular pattern in training. For example, if a model had the feature [± monosyllable] in its input representation, it might be capable of correctly generalising the regular ending to most new polysyllabic verbs, since few polysyllables are irregular.[1] To test the prediction that humans can apply the regular process as a default operation to stems with any sound, it is necessary to see how people inflect novel monosyllabic verb roots that do not resemble English regular verbs that people have had experience with. That is the primary goal of the experiments reported here. We will attempt to replicate Bybee and Moder's finding that people generalise irregular patterns to novel verbs according to their similarity to familiar ones displaying the pattern, and then, crucially, see how people extend the *regular* suffix to new verbs differing in their similarity to familiar *regular* verbs.

Several aspects of the data will be examined. First, we will simply see whether people are willing to apply the regular suffix to verbs that do not clearly resemble existing regular verbs (both verbs that closely resemble irregulars, and, more important, verbs that resemble neither irregulars nor regulars). Though it might seem like a foregone conclusion that people faced with a novel verb would apply the regular suffix to it, and that Rumelhart and McClelland's failure to do so in some circumstances is so psychologically implausible as not to be worth testing, that is not so. In some circumstances, people, too, are at a genuine loss as to how to inflect a form. Halle (1973) cites a short story written by the Russian humorist Zoshchenko, in which a nightwatchman had a difficult time ordering a set of pokers because, like most Russians, he didn't know the required genitive plural form of the word for "poker". Indeed, we already know that regular inflection is not automatically applied when a verb is strongly attracted to a highly similar irregular but is low or zero in frequency: When English speakers are asked to complete the sequence *I stride, I strode, I have* _____, most people are unhappy both with *stridden* and with *strided* (Wasow, 1981). For related reasons, neither *forwent* nor *forgoed*, and neither *forbore* nor *forbeared*, is fully acceptable (Pinker & Prince, 1988), and these judgements are easily experimentally quantified in low ratings (usually lower than the midpoint of the rating scale) for both forms (Ullman & Pinker, 1991). So the crucial question is whether strong

[1]Indeed, the only irregulars that are polysyllabic are those consisting of a prefix and a monosyllabic root, such as *understand, undergo*, and *become* (see Pinker & Prince, 1988). Therefore, if a model represents words not just as strings of sounds but as hierarchical morphological structures consisting of stems and prefixes, it could learn that all polysyllabic words other than those containing irregular roots take the regular suffix, and thus could generalise perfectly to novel polysyllabic words. Note that a representation of morphological structure (prefix + stem) is necessary to capture this law successfully, not just a representation of syllabicity: *succumb*, which rhymes with *come* and *become* but is not morphologically decomposable, is regular (Pinker & Prince, 1988).

dissimilarity to regulars (as opposed to strong similarity to irregulars) is a condition causing people to be queasy about applying the regular inflection to a novel verb.

Secondly, if people do accept and produce regularly inflected versions of any novel sound, we will seek to measure whether there is a graded effect among those forms of their relative similarity to existing regular verbs.

The principle methodological challenge in this study, especially in testing for effects of similarity, is to find verbs whose sound patterns are not contained in real verbs that human adults have previously experienced in regular past-tense forms. English-speaking adults have encountered thousands of regular verbs – for example, there are more than 4000 verbs in Francis and Kučera (1982) and more than 6000 in Webster's *Seventh Collegiate Pocket Dictionary*. Such verbs are surely found over large parts of the space of possible English sound patterns, and if it turns out that people can generalise easily to novel-sounding verbs, it may be because they had previously encountered a regular verb in the past tense that contained some of its patterns. (Of course, whether or not a new item is treated as resembling a trained one will depend heavily on the phonological representational scheme a system uses.)

In designing our stimuli, we wanted a range of degrees of similarity to existing regular verbs. Finding verbs that are most similar to regular verbs along the relevant dimensions is straightforward. Rhyme appears to be the most salient dimension of generalisation for irregular patterns (Bybee & Moder, 1983; Pinker & Prince, 1988), and so we picked verbs that rhyme with a large number of regular verbs. For example, *plip* rhymes with many monosyllabic verbs, and Francis and Kučera (1982) include the following: *chip*, *clip*, *dip*, *drip*, *flip*, *grip*, *nip*, *quip*, *rip*, *ship*, *sip*, *skip*, *slip*, *strip*, *tip*, *trip*, *whip* and *zip*. Verbs less similar to regulars can be assembled by using sequences of an initial consonant cluster and vowel, and a vowel and final consonant cluster, that appear in no English verbs, such as *smaig*. Finally, verbs with minimal similarity to existing verbs can be guaranteed by using final consonant clusters that never appear in English verbs, sequences of an initial consonant cluster and vowel that never appear in English verbs and – ensuring really unusual sounds – sequences of a vowel and final consonant cluster that violate phonological constraints on possible English verbs. For example, *ploamph* stands apart because no English verb ends in *mf* and English phonology forbids a tense vowel to precede a consonant cluster syllable-finally unless the cluster consists entirely of coronal consonants (e.g. *toast* is possible, but not *toask* or *toasp*; see Borowsky, 1989, for a discussion and an explanation). All of the verbs we assembled according to these specifications nonetheless contained a vowel common in English irregular verbs, giving subjects the option of considering irregular responses that correspond to actual vowel changes in English irregular past-tense forms. Thus, while the data of primary interest concern the willingness of subjects to apply the regular mapping to stems unlike known regulars, we are also able to measure any tendency to assign irregular mappings to them.

In addition, we included a set of verbs that range in the degree of similarity to existing irregular verbs. This allowed us to replicate the results of Bybee and Moder (1983) as well as compare the manner in which regular and irregular inflection are generalised as a function of similarity to existing regular and irregular verbs. Analogy-based models of inflection predict that similarity to known verbs is necessary for generalising both regular and irregular inflection; a rule-plus-associative storage model predicts that similarity to known verbs is necesary only for generalisation of irregular inflection. The strongest form of the rule-plus-associative memory model would maintain that predictable regularly inflected forms are never stored in memory, and hence one would see no facilitation whatsoever in inflecting new verbs that are similar to familiar ones. A weaker form would say that prior storage of regulars is possible, and thereby might offer mild analogical assistance to the generalisation of the regular inflection to similar forms, but generalisation never depends on prior storage of a similar form, so it should never fail outright, even for very unusual sounding forms.

Unfortunately, this design necessarily involves a confound: Items that are designed to have as little chance as possible of resembling English regular verbs along any dimension necessarily will have unusual sounds. Verbs with unusual sounds obviously will sound odd to native English speakers. If we were to ask people to judge the forms *plipped* and *ploamphed*, it would not be surprising if they found the latter worse, not because their past-tense inflection mechanism was impaired by not having been trained on the sounds of *ploamph*, but because the sounds of *ploamph* themselves are bad to an English speaker (because of unfamiliarity, violation of phonological rules, or both), infecting the judgement of any of its versions. The question we hope to address is not whether unusual sounds are judged to sound unusual in themselves, but whether there is any further degradation or decrement in the output when an unusual sound is put through the past-tense inflection system. Thus to measure the effects of the past-tense operation *per se*, we compare ratings of past-tense forms for novel verbs against the baseline of ratings of the uninflected stem form. This tests whether the past-tense operation degrades or weakens the integrity of the phonological representations of the input stem in the process of converting it to a past form, as a function of the familiarity of sound pattern of the stem. In the third experiment, we will measure the probability that subjects supply regular and irregular forms for novel verbs.

Finally, we investigate the patterns of generalisation under different training conditions of Rumelhart and McClelland's (1986) model, and compare it to the patterns of responses given by subjects. These simulations were designed to confirm that the model generalised both regular and irregular patterns according to similarity to items in its training set. We discuss these results in the context of more sophisticated connectionist models that have recently been proposed.

EXPERIMENT 1

In the first two experiments, we assessed people's generalisation of regular and irregular morphological patterns by having them rate the phonological goodness of artificial verb stems, and the naturalness of their regular or irregular past-tense form. Assessing judgements of well-formedness is the most appropriate dependent measure for an investigation of effects of similarity on generalisation, because the question at issue is what kinds of forms are deemed acceptable to the speaker; other psycholinguistic measures such as reaction time and speech errors are not meaningful until the prior question of what is to be counted as a "correct form" *vs* an "error" is addressed. The ratings we will elicit are the continuous equivalent of the data used in linguistic research, namely binary well-formedness or grammaticality judgements.

One possible objection to this methodology is that producing a rating is a metalinguistic act governed by conscious deliberate reasoning. However, although the conversion of an internal graded sensation of grammatical naturalness into a rating is a conscious act, the psychological processes that create that internal sensation are not: Kim et al. (1991b) document that although subjects behave lawfully in their relative willingness to apply regular and irregular inflectional patterns to different grammatical forms, they were not schooled in, nor are they aware of, the laws that governed such behaviour. A second objection might be that ratings use a coarse measurement scale in comparison to chronometric data. However, mean ratings are not necessarily any less sensitive than mean response times, because ratings are freer of sources of noise such as fluctuating attentional and sensorimotor factors.

Method

Subjects. Twenty native English speakers were paid for their participation.

Materials. We constructed a questionnaire containing 60 novel verbs, which are reproduced in Appendix 1. The verbs were used in all the experiments reported here. There were 10 novel verbs in each of six classes. The first three classes were similar to the classes used by Bybee and Moder (1983).

Items in the *prototypical pseudo-irregular class* rhymed with the prototype of clusters of similar irregular verbs. Three verbs exemplified the typical patterns of the *ing/ung* cluster (e.g. *spling*) examined by Bybee and Moder. This pattern, the prototype of which is a three-consonant cluster beginning with *s*, followed by the vowel [I], followed by a velar nasal, is exemplified in the past forms of the English verbs *cling*, *fling*, *sling*, *slink*, *sting*, *string*, *swing*, *wring*, and in the participle forms (and in many dialects, the past forms) of *ring*, *sing*, *spring*, *drink*, *shrink*, *sink*, *stink*; it is also seen in the pasts of the more peripheral members *stick*, *dig*, *win*, *spin*, *hang*, *strike*, *sneak* and the participles of *swim*, *begin* and *run*. Three items exemplified the *eed–ed* alternation found in *bleed*, *breed*, *feed*, *lead*, *read*, *speed*, *plead* and neighbouring *meet* (featurally, the vowel laxing pattern is also seen in verbs like *hide*, *bite*, *shoot*, and the [i-ɛ] pattern itself is also found in combination with suffixation in verbs like *deal*, *mean*, *creep*, *leave* and *flee*). Two items matched the prototype of verbs going to [u] in the past, consisting of a consonant cluster and a word-final [o], found prototypically in *blow*, *know*, *grow*, *throw* and peripherally in *draw*, *fly* and *slay*. The final two items matched the pattern of [ɛr/or] found in *wear*, *swear*, *tear*, *bear* and more peripherally in verbs like *get*, *tread*, *wake* and *break*.

Intermediate pseudo-irregular items were derived by taking each prototypical pseudo-irregular item and changing either its initial or its final consonant cluster (e.g. *ning*). This was the basic operation used by Bybee and Moder to derive their less prototypical items.

Distant pseudo-irregular items were derived by taking each prototypical pseudo-irregular item and changing both its initial and final consonant clusters (e.g. *nist*).

Prototypical pseudo-regular items were designed to rhyme with many regular verbs, with the constraint that they contained the vowels used in the pseudo-irregular items. This was done so that the subjects could consider irregular responses which correspond to actual vowel changes in English irregular past-tense forms. *Plip*, *brip* and *glip* rhyme with 18 monosyllabic regular verbs listed in Francis and Kučera (1982); *gloke* and *proke* rhyme with 9 (*croak*, *choke*, *joke*, *poke*, *soak*, *smoke*, *spoke*, *stoke*, *stroke*); *greem*, *pleem* and *treem* rhyme with 11 (*beam*, *cream*, *gleam*, *deem*, *scheme*, *scream*, *seem*, *steam*, *stream*, *teem*, *team*); *slace* and *nace* rhyme with 12 (*base*, *brace*, *case*, *chase*, *face*, *grace*, *lace*, *pace*, *place*, *race*, *space*, *trace*).

Intermediate pseudo-regular items were designed to be very different from existing regular and irregular verbs. They begin with consonant cluster–vowel sequences which do not exist in English verbs (e.g. [smei]), and end with vowel–consonant cluster pairs which do not exist (e.g. [eig]), e.g. *smaig* and *ploab*. It was not possible to find initial consonant cluster–vowel sequences which do not exist for the items with the vowel [I]; for these items, the final cluster is the only one that does not exist (e.g. *glinth*).

Distant pseudo-regular items were designed to be maximally different from existing verbs. They differed from intermediate pseudo-regular items in that the final consonant clusters of the latter exist in English verbs, whereas the final consonant clusters of the distant pseudo-regular verbs do not (e.g. *ploamph* and *smeerg*; see Borowsky, 1989). The initial consonant cluster–vowel sequences, like those of intermediate pseudo-regular items, except for the items with the vowel [I], also cannot be found in English verbs.

One approximate but objective common measure of the degree of average similarity of each of the classes to the relevant English vocabulary items is the mean number of monosyllabic regular and irregular verbs in Francis and Kučera (1982) that rhyme with the class members (for pseudo-irregular classes, only irregulars whose past-tense forms rhyme with the form we provided in the rating questionnaire would be included, i.e. for *spling* we would include *string* but not *bring*). The items in the three pseudo-irregular classes had, respectively, averages of 5.4, 3.3 and 0 relevant English irregular rhymes, and 6.2, 4.5 and 2.4 English regular rhymes.[2] The items in the three pseudo-regular classes had averages of 7.2, 0 and 0 regular rhymes, and no irregular rhymes.

Two kinds of past-tense forms were created for all 60 items, both pseudo-regular and pseudo-irregular. *Suffixed* past forms were created by adding *-ed* to the novel stems, following standard English orthography. *Vowel-change* past forms were created by changing the vowel according to the patterns exemplified in the prototypical pseudo-irregular set; hence *skring–skrung, frink–frunk, trisp–trusp, plip–plup, plimph–plumph* and *frilg–frulg* for the six respective classes.

The forms were spelled with regular, maximally unambiguous English spelling–sound patterns. In cases where the pronunciation of a form was conceivably ambiguous, a rhyming word was provided. Novel items that were like irregular verbs were spelled differently than the real verbs [e.g. "*cloe* (rhymes with hoe)"; instead of "*clow*").

Procedure. The items were embedded in sentences and randomly assembled in a questionnaire. The subjects saw each item as a verb stem

[2]Note that the numbers of regular rhymes may be an overestimate, because some of them, like *wing* and *square*, have noun roots rather than verb roots and therefore may not actually be relevant to the true number of attracting forms in the language, for reasons covered in the General Discussion and studied by Kim et al. (1991b).

in the sentence context *John likes to* [stem], and in either a suffixed or vowel-change past-tense form in the frame *Yesterday*, *John* [past form]. Each sentence had a 7-point rating scale next to it. Two versions of the questionnaire were created, each given to half of the subjects. Both forms contained the full set of stems, but a given subject would see only the suffixed or vowel-change form of any given stem as its past-tense form.

The subjects were presented with the following instructions, intended to minimise irrelevant normative influences on the ratings (e.g. their folk theories of "proper" English), and response biases in using the rating scale.

> This is an experiment on people's use of language. We will be asking you for two kinds of judgements about the naturalness of sentences containing novel verbs.
>
> The sentences will be presented in sets of two. The first sentence will present a novel verb. Your task on this sentence is to rate how "natural" the novel verb sounds. Your ratings should be based on the verb's pronounceability and how good it sounds. There are no right or wrong answers; we are asking for a "gut feeling" type of response. We have tried to spell the novel verbs as they sound, using standard spellings, but in some cases we thought the pronunciation might still be ambiguous, so for these we have included extra clues about how it should be pronounced, such as a rhyming word. It is the sound that is important, not the spelling; please try to ignore the verb's spelling and base your judgements just on how natural it sounds. A rating of "1" indicates that the novel verb sounds really bad or unnatural, whereas a rating of "7" indicates that the verb sounds perfectly good and natural. Circle the other numbers for judgements that fall in between.

Two examples were then presented, *John likes to blicket* and *John likes to bvonfrt*: We said that if the subjects' judgements were like ours they would rate them as "7" and "1" respectively (these extreme examples were chosen so as not to give the subjects any biases in rating the experimental items). The subjects were encouraged to take as much time as they needed and were asked not to look ahead or go back to other sentences in the questionnaire. Finally, they were reminded that "there are no 'right' or 'wrong' answers; we are interested in your personal judgements".

The instructions for the second sentence were as follows:

> The second sentence in each set presents the novel verb in a past-tense form. Your task for this sentence is to rate how good or natural the verb is as the past tense of the verb presented in the previous sentence. We now want you to ignore how good the verb in the first sentence (present-tense form) sounds, but concentrate on how good the form in the second sentence is AS THE PAST-TENSE VERSION OF THE VERB IN THE FIRST SENTENCE. That is, this time we are not interested in pronounceability itself, but

> in the naturalness of the sound of the verb as the past tense of the new verb. For example, the word *drunt* is perfectly pronounceable, but it is a terrible-sounding past-tense version of the verb "to drobe". Again, it is the sound as the past-tense form we want you to judge, NOT spelling.

As before, examples of typical "1" and "7" items were given (*dance–danced, sing–singed*), chosen so as not to provide information biasing responses for the experimental items. The subjects were again told that there were no "right" or "wrong" answers.

Results and Discussion

The mean ratings of the different word classes are shown in Fig. 1. Looking first at the relative strength of the vowel-change (irregular) and suffixed (regular) patterns overall, we find that the regular pattern wins out even in the circumstances most favourable to irregular generalisation. In all three pseudo-irregular classes, the suffixed past form was given higher average ratings than the vowel-change form. Of the 30 individual pseudo-irregular verbs, only one, *fring* (in the intermediate class), showed a significantly higher rating for the vowel-change than the suffixed form at a two-tailed *P* of 0.05. An additional four items had non-significantly higher ratings for the vowel-change than the suffixed form: *preed, quare* and *spling* from the prototypical pseudo-irregular class, and *ning* from the intermediate pseudo-irregular class. For the remaining 25 pseudo-irregular verbs, the suffixed form was rated higher, 16 of them significantly (3 out of the 10 prototypical pseudo-irregular verbs, 5 of the 10 intermediate pseudo-irregular verbs and 8 of the 10 distant pseudo-irregular verbs).

These results suggest that despite the fact that irregular patterns can occasionally be generalised to new forms, regular suffixation is the more quantitatively powerful generalisation mechanism. (The other experiments we report confirm this difference.) The greater strength of the regular process in competition over irregular-sounding stems is consistent with one of the major trends in the history of English, documented by Bybee (1985): When the frequency of an irregular verb declines, it becomes regular. Since most irregulars in the history of the language belonged to clusters of similar forms, they could only have been regularised if the strength of regular suffixation substantially exceeded that of irregular analogy, because otherwise the attraction of an irregular verb's family members would have sufficed to hold it in the cluster, no matter how low its frequency slipped.

To assess generalisation as a function of distance from an irregular cluster, three one-way repeated-measures ANOVAs were performed on the ratings of pseudo-irregular verbs, one each on the ratings of stem, vowel-change and suffixed forms, with distance from English irregular clusters as the independent variable. Not surprisingly, there is no difference in the phonological goodness of stem forms at different distances from an irregular cluster [by subjects: $F_1(2,38) < 1$; by items: $F_2(2,27) < 1$]. Replicating Bybee and Moder (1983), we find a significant decline in the goodness of vowel-change forms as a function of distance from an irregular cluster [$F_1(2,38) = 17.11$, $P < 0.0001$; $F_2(2,27) = 7.19$, $P < 0.005$]. There was also a very small and non-monotonic difference among the mean goodness ratings of the suffixed forms, which was marginally significant in the subject analysis [$F_1(2,38) = 3.22$, $P < 0.051$; $F_2(2,27) < 1$]. The pattern is not completely interpretable but may reflect a graded version of

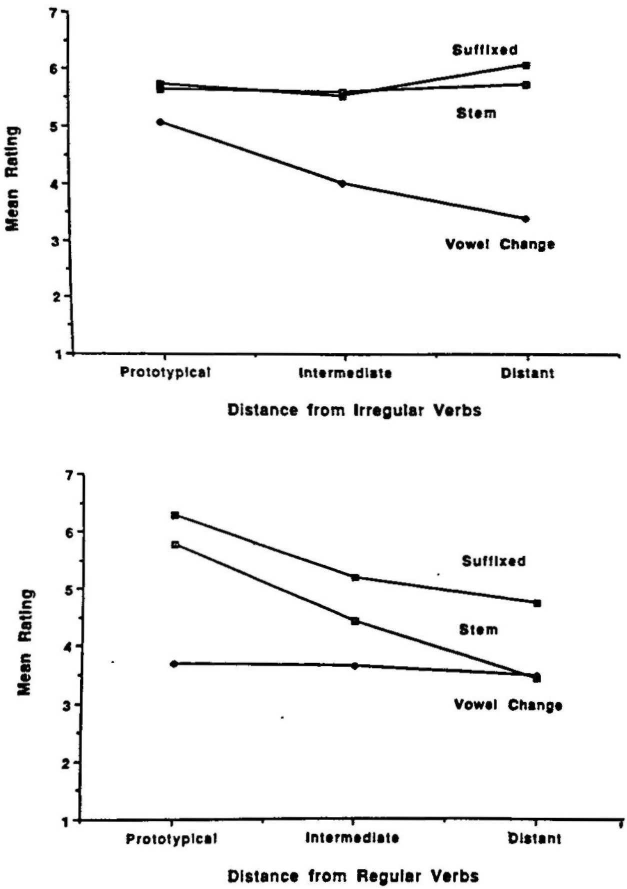

FIG. 1 Mean ratings of naturalness of stems and past-tense forms in Experiment 1. (Top) Pseudo-irregular verbs; (bottom) pseudo-regular verbs.

the blocking of regularisation by irregular-sounding roots: The more a given stem evokes existing irregular past forms, the more it inhibits the derivation of a regular (suffixed) form. Note that this is unlikely to be due to increasing attraction to regular forms in the peripheral class, as there are fewer English regular verbs rhyming with the items in this class.

For pseudo-regular classes, suffixed forms for all classes were given high mean ratings (4.8 for the worst class, which is close to the 5.1 rating for the *best* class of vowel-change forms of pseudo-irregular verbs). Vowel-change forms, and stems, were rated lower for all three classes. Both the stem and suffixed forms showed a decrease in goodness as a function of distance from regular clusters [for stems: $F_1(2,38) = 56.16$, $P < 0.0001$; $F_2(2,27) = 65.57$, $P < 0.0001$; for suffixed forms: $F_1(2,38) = 20.55$, $P < 0.0001$; $F_2(2,27) = 13.45$, $P < 0.0001$]. Ratings of the vowel-change form did not vary as a function of distance from a regular cluster ($F < 1$). The ratings of these forms, however, were extremely low.

As mentioned, a decline in the goodness of the suffixed form with distance from a cluster of existing verbs could reflect either a degradation of the past-tense formation process for unusual sounding stems, or direct effects of the extreme phonological unusualness of the intermediate and distant items on the judgements, regardless of the past-tense process. Though the instructions were intended to encourage the subjects to rate all items *as* the past-tense forms of a given verb, it is possible that they did not block phonological factors out of their judgements, especially since the instructions contained numerous (in hindsight, unfortunate) uses of the word "sound", such as "the naturalness of the sound of the verb as the past tense of the new verb", "terrible-sounding", and "sounds bad" and "sounds good" at the endpoints of every rating scale. "Sounds" is ambiguous between phonological goodness and grammatical naturalness and, though the latter was intended, the former may have played a role as well.

To see if the past-tense mechanism actually works more poorly for items with more unusual sounds, we compared ratings of the past-tense forms with those of the stem itself. As Fig. 1 shows, the pseudo-regular stem forms themselves (not subject to any internal transformation or analysis) decline in goodness from prototypical to intermediate to distant. But whatever process handles the past tense clearly is not further degrading the integrity of the internal representation of the word when analysing it as a past form: The decline of the goodness of the regular past form with distance is *less* than the corresponding decline for the stem form itself. (Presumably, the presence of a familiar verbal inflection made the distant items seem more English-like than they are as pure, and somewhat bizarre, morphemes.) This is in direct contrast to vowel-change forms of the pseudo-irregular items, where the goodness of the past form declines much more than the stem as a function of distance. The difference between pseudo-irregular and pseudo-regular classes can easily be seen when the rating of the relevant past form is subtracted from the rating of the corresponding stem, as is shown in Fig. 2. It is confirmed in a 3 × 2 × 2 analysis of variance with factors pseudo-irregular/pseudo-regular class, prototypical/intermediate/distant, and stem/past form (stem form for both classes, versus suffixed form for pseudo-regular classes and vowel-change form for pseudo-irregular classes). The three-way interaction is significant [$F(2,38) = 20.511$, $P < 0.001$]. This analysis must be interpreted with caution, however, because the distance metrics for pseudo-irregular and pseudo-regular classes are not directly comparable. A test that would minimise this problem is a one-way analysis of covariance on the past-tense ratings using each verb's mean stem rating as the covariate; this subtracts out any linear effect of stem ratings on past-tense ratings, and is independent of differences in the subjects' internal scale units or floor. An unconfounded effect of distance from irregular clusters was found for the ratings of vowel-change past-tense forms of pseudo-irregular verbs [$F(2,26) = 7.61$, $P < 0.005$], but no unconfounded effect of distance from regular clusters was found for the ratings of suffixed past-tense forms of pseudo-regular verbs [$F(2,26) < 1$]. Thus for regular pasts, unlike irregu-

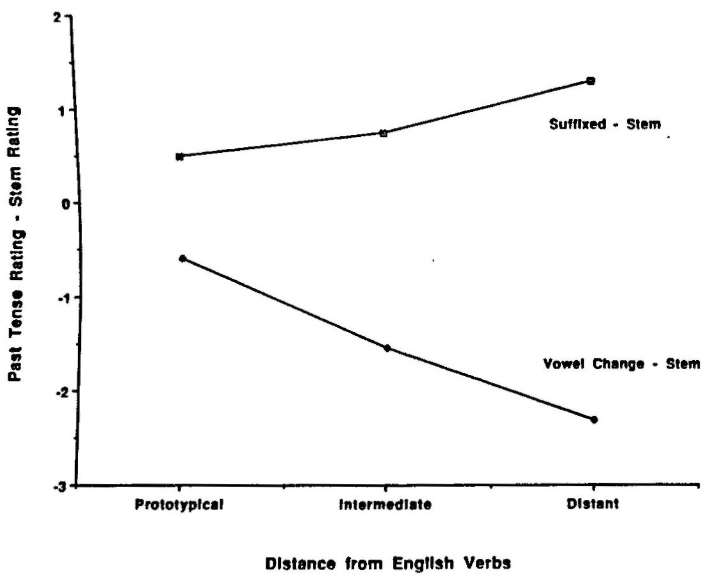

FIG. 2 Experiment 1. Effects of similarity to known verbs on the generalisation of regular (suffixed) and irregular (vowel-change) patterns to novel stems, holding stem naturalness constant. Each line corresponds to the difference between ratings of the past-tense form and ratings of the stem form.

lars, there is no statistical basis for attributing the decrease found in the ratings in the experiment to anything other than the decline in the goodness of the sound of the stem itself.

The results of the current experiment are consistent with the findings of Bybee and Moder (1983) that subjects' preference for past forms inflected like irregular verbs decreases as novel forms become less similar to clusters of those irregular verbs. In the present experiment, past-tense forms based on irregulars decreased in goodness as the stems became less similar to clusters of English irregular verbs, with irregularised dissimilar forms given quite low ratings. This is easy to account for if irregularly inflected verbs are stored in an associative memory. The results of the experiment suggest that the regular inflection process might be a different kind of phenomenon, at least quantitatively and perhaps qualitatively. All classes of suffixed past-tense forms, whether similar or dissimilar to existing regulars, elicited high ratings (well into the top half of the scale). Although there was a measurable decline in the goodness of the suffixed forms as the stem was increasingly unlike familiar English verbs, one cannot conclude that this reflects an effect of familiarity on the past-tense formation or analysis process, because pure phonological familiarity effects measured on the verb stem account for the entire decline, with no added effect of past-tense formation *per se* visible in the data. The data suggest that regular inflections may be produced by the application of some psychological process that is insensitive to the phonological properties of the stem to which it is applied (including the familiarity of those properties), with the result that the regularly inflected form simply inherits the familiarity (or lack thereof) of the stem.

EXPERIMENT 2

This experiment was conducted to determine the generality of the results of Experiment 1. In Experiment 1, the subjects were required to judge the phonological goodness of a stem form, and then immediately to judge the goodness of the same item converted to a past-tense form, on every trial. It is possible that the subjects may have perseverated in judging phonological goodness when rating past-tense forms more than they would if they didn't have to keep switching back and forth. In the present experiment, the subjects rated all past-tense forms first, and then rated all the stems. In addition, the subjects rated how "likely" a given item was as the past-tense form of a given verb, rather than how good it "sounds" as the past-tense form, to emphasise that we were not interested in the sound of the past-tense forms *per se*.

Method

Subjects. Twenty-two native English speakers were paid for their participation.

Materials. The materials were the same as in Experiment 1, except that the questionnaire was divided into two parts, the first asking for ratings of past-tense forms, the second asking for ratings of the stems. The design was the same as that in Experiment 1, and involved two versions of the questionnaire, each given to half of the subjects, differing in whether a given item was presented in the vowel-change or suffixed past-tense form.

Procedure. The instructions for the first part of the questionnaire asked the subjects to rate:

> ... the likelihood of the word **as the past tense of the given verb**. For example, most people would probably feel that the word *drunt* is not likely to be the past-tense version of the verb *to drobe*. It is only the likelihood that the word is **the past-tense form** we want you to judge. A rating of "1" would indicate that the word is a very unlikely candidate for **the past tense of the given verb**, whereas a rating of "7" would indicate that the word is a very likely candidate for **the past tense of the given verb**.

The instructions for the second part were identical to those for Experiment 1, asking for judgements of how good the verbs sounded. Each item was of the form *John likes to* [stem].

Results and Discussion

The results are shown in Fig. 3, and are similar in most respects to those of Experiment 1 (Appendix 1 presents the mean ratings of stem and past-tense forms for each item, averaged across Experiments 1 and 2). Among the pseudo-irregular items, the suffixed forms were on average rated slightly (though not significantly) higher than the vowel-change forms even for the prototypical items. Among the 10 prototypical items, only *queed* was rated significantly higher in its vowel-change form, and *cleed*, *froe*, *skring* and *spling* were rated non-significantly higher; among the 10 intermediate items, *ning* was rated significantly higher as a vowel-change form and *fring* non-significantly higher. Of the 23 remaining verbs, the suffixed form was rated significantly higher for 3 of the 5 in the prototypical class, 5 of the 8 in the intermediate class and 9 of the 10 in the distant class. Thus again we find that pseudo-irregular forms, even those highly similar to existing English irregular clusters, only rarely support generalisations to

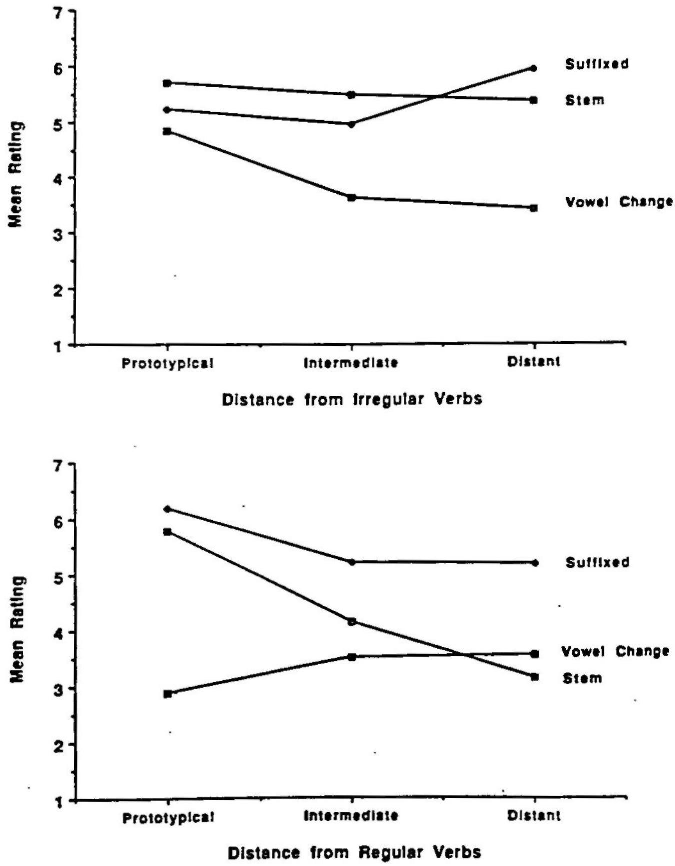

FIG. 3 Mean ratings of naturalness of stems and past-tense forms in Experiment 2. (Top) Pseudo-irregular verbs; (bottom) pseudo-regular verbs.

irregular (vowel-change) inflected versions strong enough to surpass the regular suffixed version significantly.

As in Experiment 1, there was no difference in the phonological goodness of pseudo-irregular stem forms [$F_1(2,42) = 2.12$, NS; $F_2(2,27) < 1$] as a function of their distance from an irregular cluster. And as in Experiment 1 and Bybee and Moder (1983), there was a significant decline in the goodness of vowel-change forms as a function of distance [$F_1(2,42) = 19.16$, $P < 0.0001$; $F_2(2,27) = 6.39$, $P < 0.005$]. Finally, we replicated a non-monotonic increase for non-prototypical suffixed forms of pseudo-irregular verbs [$F_1(2,42) = 10.18$, $P < 0.0001$; $F_2(2,27) = 3.33$, $P < 0.051$], perhaps reflecting partial blocking by existing similar English irregular forms. (As noted in Experiment 1, it cannot be due to an

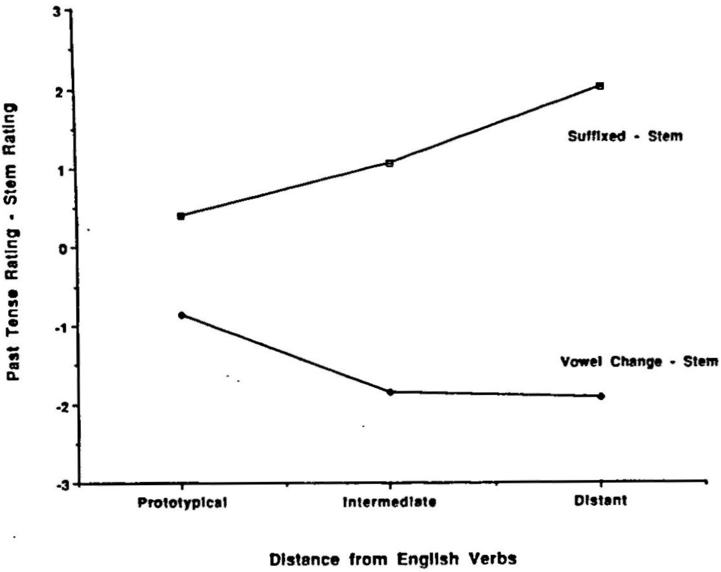

FIG. 4 Experiment 2. Effects of similarity to known verbs on the generalisation of regular (suffixed) and irregular (vowel-change) patterns to novel stems, holding stem naturalness constant. Each line corresponds to the difference between ratings of the past-tense forms and ratings of the stem form.

increasing tendency to analogise items to English regular verbs, because the number of rhyming English regulars *decreases* with distance from the irregular prototype.)

Again pseudo-regular items were given high mean ratings in all conditions in their suffixed forms. For pseudo-regular items, both the stem and suffixed past-tense forms showed a decrease in goodness [for stems: $F_1(2,42) = 120.24$, $P < 0.0001$; $F_2(2,27) = 68.09$, $P < 0.0001$; for suffixed pasts: $F_1(2,42) = 12.95$, $P < 0.0001$; $F_2(2,27) = 6.78$, $P < 0.005$]. Interestingly, there is no difference between ratings of the suffixed pasts of intermediate and distant items, despite a large difference in the ratings of the stem. Ratings of the vowel-change versions of these items increased as a function of distance from a regular cluster in the subjects analysis [$F_1(2,42) = 5.26$, $P < 0.01$], but not in the items analysis [$F_2(2,27) = 2.16$, NS].

Once again, we sought to determine whether the effects of distance from regular and irregular prototypes affected the goodness of the corresponding past-tense forms above and beyond the goodness of the stems. In Fig. 4, past-tense ratings are subtracted from stem ratings; we see that the pseudo-regular and pseudo-irregular past forms show opposite effects of distance from a prototype in comparison with their stem forms. This is confirmed in a significant three-way interaction between pseudo-regular/pseudo-irregular class, prototypical/intermediate/distant, and stem/past form (stem form for both classes, versus suffixed past form for pseudo-regular classes and vowel-change past form for pseudo-irregular classes) [$F(2,42) = 28.02$, $P < 0.001$]. Similarly, when stem ratings of different verbs were separated and entered as the covariate in analyses of covariance, an effect of distance from irregular clusters was found for the ratings of vowel-change past-tense forms of pseudo-irregular verbs [$F(2,26) = 7.05$, $P < 0.004$], but no effect of distance from regular clusters was found for the ratings of vowel-change past-tense forms of pseudo-regular verbs [$F(2,26) = 1.13$, NS]. This confirms the hypothesis that the goodness of vowel-change past-tense forms declines as a function of distance from irregular clusters and that this decline cannot be attributed to phonological factors, but that the goodness of the suffixed past-tense forms does not decline as a function of distance from known suffixed forms, holding constant the naturalness of the stem in both cases.

The only difference between this experiment and Experiment 1 was that ratings of vowel-change pseudo-regular forms increased as a function of distance from regular clusters. We suspect that the increase may reflect the fact that regularly inflected forms in this part of phonological space, because they involve the addition of *-ed* to an already unusual sounding stem, are particularly unpalatable as phonological objects. If subjects mentally compare a given vowel-change form with the corresponding regularly suffixed version, they might rate the vowel-change form higher when the suffixed form is phonologically awkward. As part of a related investigation, we gathered norms from 24 different subjects in which they rated the phonological goodness of all 180 forms as phonetically spelled nouns (these norms are reproduced in Appendix 1 and, averaged within categories, in Table 1). Stem ratings behaved like those in Experiments 1

TABLE 1
Mean Ratings of the Phonological Naturalness of All Word Forms Presented in Experiments 1 and 2

	Stem	*Vowel-change*	*Suffixed*
Prototypical pseudo-irregular	5.6	5.3	4.0
Intermediate pseudo-irregular	5.2	5.2	4.5
Distant pseudo-irregular	5.5	5.6	3.4
Prototypical pseudo-regular	5.7	5.2	3.9
Intermediate pseudo-regular	4.3	4.8	2.5
Distant pseudo-regular	3.4	3.9	2.5

and 2, but note that the ratings of the suffixed forms were considerably lower and, as expected, suffixed versions of the intermediate and distant pseudo-regular items were extremely low.

EXPERIMENT 3

This experiment sought to determine the likelihood of people actually producing regular and irregular past-tense forms, rather than only rating them. The use of a production task provided an opportunity to replicate the results of the previous experiments using a task that is more natural than providing ratings, and that reflects more of the language system, including any internal criterion used to translate continuous strength or goodness values into a binary decision to use one or another form. It also allows us to assess whether people have any tendency to modify unfamiliar stems in ways other than those based on analogies with existing irregular verbs. In addition, we introduced several procedural changes designed to reduce the extent to which subjects' ratings might be based on the phonological properties of the items.

In Experiments 1 and 2, the novel forms were all presented as verbs; the only property that the forms shared with real verbs was the syntactic frame in which they were presented. Each trial had a verb in it, and it is possible that the subjects were treating the novel forms as phonological strings rather than verbs. Clearly, this could not explain the full range of results, but aspects of the task may have led to phonological factors being weighted more highly than would be the case in ordinary speech. The current experiment used the novel verbs used in Experiments 1 and 2, as well as novel nouns and adjectives. Each novel word was given a definition and the subjects were asked to use the words to fill in blanks in test sentences. We hoped that this would encourage the subjects to treat the items like real words rather than mere sounds. Finally, we encouraged the subjects to provide more than one past-tense form, by having them make a second pass through the questionnaire and inviting them to supply other forms that came to mind.

Method

Subjects. Twenty-four native English speakers were paid for their participation.

Materials. Each questionnaire included the 60 novel verbs used in Experiments 1 and 2 as well as 30 monosyllabic nouns and 30 adjectives with the same general properties as the verb stems. A definition was given for each of the 120 novel forms. There were six versions of the questionnaire, each presented to four subjects. The questionnaires differed in terms of which meanings were paired with which stems; across the experiment, each meaning appeared once with an item from each of the six classes.

At the top of each page, the subjects saw 6 of the 120 novel items, along with their definitions. This was followed by six sentences with blanks in them where one of the novel forms belonged. Each sentence was followed by a 7-point rating scale which had 1 labelled as "very unlikely" and 7 as "very likely". For example, one of the six novel items at the top of the page was presented as "**cleef**: to slide from side to side". Farther down the page, one of the six test items was "The ladder was built to move sideways so that it _____ as the librarian needed to reach far away books", followed by the rating scale.

Procedure. The subjects were asked to read the words and their definitions at the top of each page carefully and then to:

> . . . fill in the blank with a form of the proper novel word. It is important that the form matches the context of the sentence. For example, in the example given above, it is important that you write what you think the plural form of *wug* will be since the context requires a plural. If instead of a novel word the real word *dog* had been used, then one would fill in the form *dog*s and not *dog*; if the word were *man*, one would fill in the blank with the form *men*. After you have filled in the blank, we would like you to rate how likely it is that the form you filled in is the correct one for that word in that context. By "correct" form we mean the one that other native English speakers would naturally and intuitively use; we aren't referring to the kind of "proper" English that gets taught in grammar lessons or style manuals.

In sentences where the novel word was a noun, the subjects had to fill in a singular or plural form of the word; in sentences containing adjectives, either a stem form or a comparative form was appropriate; in sentences containing verbs, stems or past-tense forms were called for. Which of these forms was required was clear from the sentential context. The instructions emphasised that the subjects should rate each form independently:

> When rating the likelihood of a given form please rate each form independently; don't worry about how many times you use each point on the rating scale. That is, if all the forms you fill in seem very likely to you, it is perfectly all right to circle only 7s. Conversely, if you feel confident about some of the forms but not others, you could have some ratings at 7, some at 1, and some in between, in whatever proportion happens to come out.

Once the subjects had provided and rated one form for each item, they were asked to go through the questionnaire a second time. They were told that some words took two inflected forms, such as *dreamed* and *dreamt* and *fish* and *fishes*, and were given instructions that we hoped would encourage (but not coerce) them to provide such alternatives if they thought of any:

> . . . if there are cases where there are other forms that you feel would be likely forms of the word please write them in with a rating for the alternative form. Please do NOT feel challenged or obliged to think of alternatives. On the other hand, if there are alternatives that occur to you and that you feel are likely in that context, don't hesitate to list them. You could list no alternatives, or alternatives for every item, or alternatives for some items; whatever genuinely reflects your intuitions.

In the final section of the questionnaire, the subjects rated all 60 stem forms.

Results and Discussion

All responses, both first and subsequent choices, were counted. Responses were coded as "suffixed" if (1) an *-ed* was added to the stem, (2) a *t* was added when the last consonant was voiceless and (3) if a *d* was added when the stem ended in an *e*. All other responses (mainly vowel-change, but also some forms where suffixation accompanied a vowel change as in the existing *sleep-slept*, and any consonant changes) were coded as "non-suffixed".[3] Thus it is possible for there to be more than one non-suffixed response for a given item for a given subject.

The overall preference for suffixed past-tense forms for novel stems of all kinds is seen even more strongly in these production data than in the ratings data from the preceding experiments. For all six classes, there were significantly more suffixed than non-suffixed responses (all $Ps < 0.05$), and only 4 of the 60 items (*preed*, *queed*, *spling*, *cleef*) elicited at least as many non-suffixed as suffixed responses.

Consistent with the results of Experiments 1 and 2, the subjects produced significantly fewer non-suffixed past-tense forms as a function of distance from an irregular cluster. The average numbers of non-suffixed forms produced for the prototypical, intermediate and distant pseudo-irregular classes were 5.9, 4.6 and 3.7 out of 10 [$F_1(2,44) = 7.20$, $P <$

[3]Note that this scoring procedure biases responses towards being scored as having non-suffixed changes, because any spelling mistakes made in copying the form's stem will lead to the response as being scored as having a non-regular pattern. This could influence scores for pseudo-regular intermediate and distant verbs, which are unlike any existing verbs and often have unfamiliar sound and spelling patterns.

0.005; $F_2(2,27) = 9.20$, $P < 0.001$]. The subjects' ratings of their own non-suffixed forms show a similar trend, with the mean ratings for the prototypical, intermediate and distant classes being 4.7, 4.2 and 4.2. A subjects analysis was not performed on the ratings data because not all of the subjects produced non-suffixed forms for verbs in each class. The items analysis was not significant [$F_2(2,27) = 2.06$, NS]. The mean ratings of the stem forms were 5.6, 5.1 and 5.3 for the three respective classes.

For pseudo-irregular items, the number of suffixed responses increased as a function of distance from irregular clusters, with means of 0.8, 0.9 and 0.9 responses per item for the three classes [$F_1(2,46) = 7.92$, $P <$ 0.001; $F_2(2,27) = 4.15$, $P < 0.05$]. This finding may reflect morphological blocking: The more likely subjects are to think of irregular-like past-tense forms, the less likely they are to produce a regular past-tense form (this is not a mathematical necessity, as the subjects were allowed to write in more than one form). A non-monotonic and therefore difficult-to-interpret increase was found for the subjects' ratings of their own suffixed forms [$F_1(2,46) = 3.84$, $P < 0.05$]; the effect was not significant in the items analysis, however [$F_2(2,27) < 1$].

For the pseudo-regular classes, suffixed forms were readily supplied in all cases: The subjects gave an average of 9.5, 9.2 and 8.9 suffixed responses out of a possible 10 for the prototypical, intermediate and distant items. The slight decrease in the number of suffixed responses produced as a function of similarity was not significant [$F_1(2,46) = 1.96$, NS; $F_2(2,27) =$ 2.51, NS]. The subjects' ratings for these responses did not monotonically decline from prototypical to intermediate to distant: 5.5, 5.2 and 5.3, respectively, a difference that was only marginally significant [$F_1(2,46) =$ 2.79, $P < 0.07$; $F_2(2,27) = 3.23$, $P < 0.06$]. The lack of a significant decrease in ratings cannot be explained away as a floor effect, whereby the subjects only produced overt responses that met some minimal internal rating, because the subjects gave lower ratings for responses produced for other classes of novel verbs. The mean ratings of the stems were 5.7, 4.5 and 3.8, respectively.

The number of non-suffixed responses provided for pseudo-regular items increased with distance from regular clusters, with means of 0.2, 0.3 and 0.3 responses per item for the three classes [$F_1(2,46) = 3.24$, $P <$ 0.05]; the increase was not significant in the items analysis, however [$F_2(2,27) < 1$]. Note that the number of non-suffixed forms produced was much lower than the number of suffixed forms, even though the subjects were capable of providing any number of non-suffixed forms but by definition could only provide a single suffixed form. The class with the greatest number of non-suffixed responses had an average of 3.7 such responses out of 10, compared with 8.9 suffixed responses for the same class.

The subjects produced non-suffixed responses at a rate of 0.24 per item among the prototypical pseudo-regular items, 0.34 per item among the intermediates and (as mentioned) 0.37 per item among the distant ones. The majority of these responses (75, 72 and 90%, respectively) involved apparent analogies to irregular past forms. These included no-change responses, changes of vowels to those found in irregular verbs (e.g. *plip–ploap*, *smeeg–smeg*, *smairg–smorg*) and a changed vowel (virtually always lax) accompanied by suffixation of *t* or *-ed* (e.g. *greem–gremt*, *smeeb–smebbed*, *smeelth–smelthed*), a pattern seen in about 16 English irregular verbs (e.g. *leave–left*, *creep–crept*, *lose–lost*, *sleep–slept*, *tell–told*, etc.). A small number involved the incorrect choice of suffix, as in *brip–brips*, *smeeg–smeegs*, even *krilg–krilgated* and *trilb–trilbous*. Finally, for the distant pseudo-regular items, there were a few responses (six) where the subjects both added a regular suffix and changed the stem's final consonant cluster, such as *trilb–trilved*, *smeelth–smeethed* and *smeenth–smeented*. No responses involved radical changes of the stem structure, such as changes of initial consonants, syllable structure or addition of segments. The source of the occasional stem-final distortions we recorded is uncertain. They might be inspired by the 18 or so existing *t-* or *d-*final irregular verbs with final consonant cluster changes such as *bend–bent*, *teach–taught*, *stand–stood*, *make–made*, *have–had*, *lose–lost* and *leave–left*. Alternatively, the stem changes may also have been made to render the form more consistent with general English sound patterns, especially in classes where our norms (see Table 1 and Appendix 1) suggest that the consonant clusters produced by adding *-ed* result in phonological awkwardness; forms like *smeelthed* in particular seemed to have triggered occasional searches for more comfortable pronunciations, just as they were given particularly low ratings in Experiments 1 and 2 and in the phonological goodness norms.[4]

The ratings for the non-suffixed forms increased from 3.7 to 3.9 to 4.9 for the prototypical, intermediate and distant classes, an increase also seen in Experiment 2. An ANOVA was not carried out on the ratings data

[4]We suspect that that the analogical creation of irregular patterns might also be highly susceptible to metalinguistic factors. The ubiquity of puns, jokes and word-play depending on irregular patterns (e.g. *I got scrod at Legal Seafood*; see Lederer, 1989, for many examples) suggests that people can be quite conscious of irregular patterns and whether they produce them might reflect in part how they construe the task; in particular, whether it is treated as a test of their creative ingenuity, and how they set their internal criterion for which low-strength analogies and associations that pop into mind are worthy of putting down on paper. The fact that we provided many obvious rhymes to existing clusters of rhyming verbs for the irregulars (e.g. *spling*) may have made sound-based strategies even more salient to the subjects. In contrast, the everyday addition of the regular suffix to novel patterns like *Yeltsin out–Gorbachev'd Gorbachev*, does not appear to be a self-conscious performance.

because a number of subjects did not produce any non-suffixed forms for one or more classes.

The results of the current study replicate Bybee and Moder's (1983) finding that subjects are more likely to produce irregular-like non-suffixed past-tense forms for novel verbs when the novel verbs are similar to the prototype of an irregular cluster. The results also show that the similarity of a novel form to familiar regular forms does not significantly affect the likelihood of subjects' producing a regular suffixed form. The results of Experiments 1–3 support a mechanism that generalises irregular inflections on the basis of close analogies to existing forms, but generalises regular inflections on the basis of a process that is not dependent for its success on how similar the phonological properties of the stem are to previously encountered verbs undergoing that morphological alternation. A rule for regulars coupled to an associative memory for irregulars is one such mechanism; we discuss it and alternatives in the General Discussion.

SIMULATIONS

Analogy-based models generalise from familiar items to novel items on the basis of similarity. In parallel distributed processing (PDP) models, items are represented as activations over a set of representational units; the similarity of two items in such a model at the input level is equivalent to the number of representational units the two items share, and graded generalisation by similarity is thus automatic. The simulations investigated the patterns of generalisation in a particular analogy-based model, Rumelhart and McClelland's (1986) model of past-tense inflection, under different training conditions. The model's performance was compared with the patterns of responses found in Experiments 1–3. Pinker and Prince (1988), in their analysis of the model, noted that it had considerable trouble in producing regular past-tense forms for certain novel verbs, even though those verbs were not obviously subject to interference from irregulars. They attributed this difficulty to the fact that the model generalises by similarity in both irregular and regular cases, and hence has inherent problems in generalising regular processes to items dissimilar from the regular items it had previously been trained on. By manipulating the similarity of novel forms to familiar forms and measuring the trained model's response to such forms, we can test this diagnosis, and by comparing the model's behaviour to humans, one can assess whether any such generalisation failures are psychologically realistic.

The Model

We used a portable implementation of Rumelhart and McClelland's (1986) model, devised by Clayton McMillan of the University of Colorado, Boulder. The model consists of a pattern associator with 460 input and output units and no hidden units, with weights adjusted during learning using the perceptron convergence procedure (Rosenblatt, 1962). The output units are turned on probabilistically, with the probability of a given output turning on increasing with the difference between the net activation of the unit and its threshold. The pattern associator itself is fed by a fixed encoding network that turns phonological representations (ordered strings of phonemes) into Wickelfeature representations (unordered sets of Wickelfeatures, each an ordered triple of phonological distinctive features); this is the sole representation of a word used in the pattern associator. The pattern associator feeds into a decoding/binding network, designed to turn the output vector (a pattern of activation over unordered Wickelfeature units) back into ordered strings of phonemes.

Simulation 1

To ascertain the model's ability to generalise past-tense mappings, we trained it on the set of 422 verbs used by Rumelhart and McClelland, of which 341 were regular and 81 irregular – these are the 422 most frequent *-ing* forms from Kučera and Francis' (1967) ranking of verbs by frequency. The model was trained for 200 epochs. The response strength, a measure of the number of output nodes correctly turned on or off,[5] averaging over all verbs, exceeded 0.95 for the last 10 epochs, indicating that the model had successfully learned the mappings in the training corpus. The model was then tested on the prototypical, intermediate and distant pseudo-irregular and pseudo-regular novel verbs used in Experiments 1–3. To get a rough confirmation that the assumed English clusters were representatively sampled in the model's training set, we counted the number of verbs in the set that rhymed with members of each of the six classes of novel verbs. Indeed, pseudo-irregular items had 16, 18 and 4 irregular rhymes in the training set of items in the prototypical, intermediate and distant classes, respectively (and 11, 8 and 7 regular rhymes). Pseudo-regular items were represented by 9, 0 and 0 regular rhymes, respectively (and no irregular rhymes). This is consistent with our assumed direction of prototypicality across classes.

The response measure used to evaluate the model's generalisations was the output of the decoding/binding network, consisting of one or more whole words reconstructed from the output Wickelfeature vector. These outputs can be compared directly to the forms produced by human subjects in Experiment 3. Figure 5 compares the mean number of responses per item provided by the subjects in Experiment 3 and the simulation, divided

[5]The response strength of an item is given by: (hits / hits+misses) – (false alarms / false alarms+correct rejections).

into suffixed responses and vowel-change responses. Because the simulation occasionally outputted bizarre strings that were not clearly related to the stems, we focused only on its vowel-change responses, not on all its non-suffixed responses. In order to compare these responses to those of humans, we also focused only on the vowel-change responses from the subjects in Experiment 3. Thus the data in Fig. 5 represent a subset of the non-suffixed responses whose means were reported in the Results section of Experiment 3.

Among the pseudo-irregular items, the network produced a few patterns analogous to those seen in similar existing English irregulars, in particular,

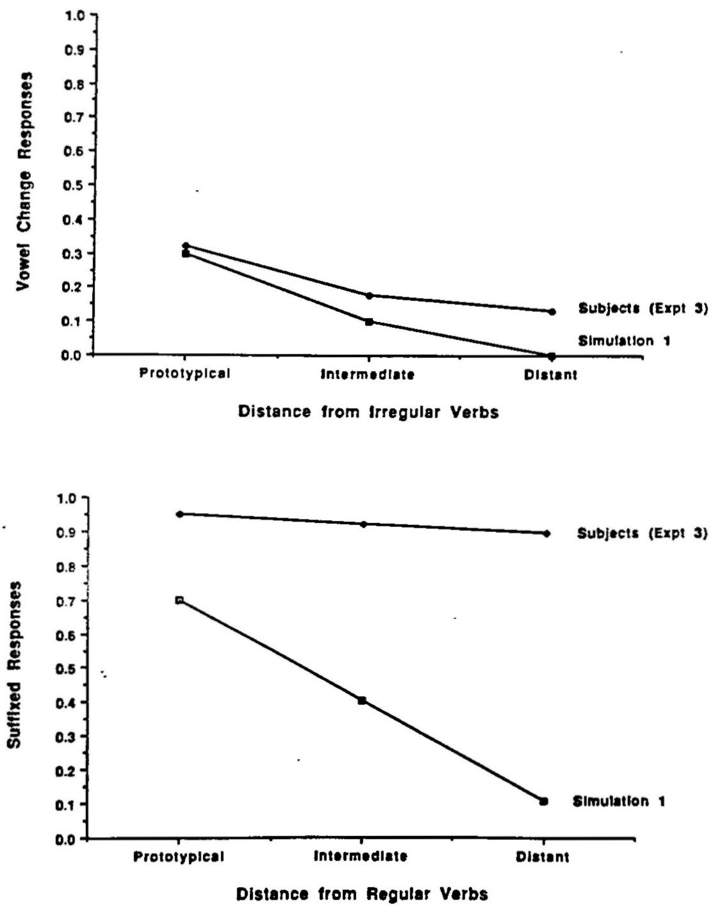

FIG. 5 Mean number of past-tense forms per stem provided by subjects in Experiment 3 and by Rumelhart and McClelland's model in the first simulation. (Top) Pseudo-irregular verbs; (bottom) pseudo-regular verbs.

vowel-change and no-change forms: three among the prototypical items (*kleed–kled*, *preed–pred*, *kweed–kweed*), one among the intermediate items (*frink–frunk*) and none among the distant items. The decline in vowel-change responses with distance from a prototype is to be expected, and is not incompatible with what we saw among our subjects.

Among the pseudo-regular items, the network produced suffixed past-tense forms for 7, 4 and 1 of the 10 items in the prototypical, intermediate and distant sets. This compares with an average of 9.5, 9.3 and 8.9 out of 10 regular forms produced by the subjects in Experiment 3. Clearly, the model had much more trouble than our subjects in generalising the regular pattern to items that are unlike the items on which it is trained. The model was also more creative than our subjects in providing alternatives for the regular items (though when it did provide a regular form, it was usually its first choice). The model produced between 3 and 10 forms for each verb, a few based on irregular analogies (e.g. *greem–grame*, *proke–prokt*, and no-change forms), some involving assimilation to particular English verbs (e.g. *brilth–prevailed*, *ploag–pleaded*, *proke–trusted*, *krilg–brewed*) and some involving a slight stem change (e.g. *brilth–brilt*, *smabe–smaned*). However, most of the non-suffixed responses consisted of unsystematic blends and distortions (e.g. *slace–fraced*, *smeeb–imin*, *ploanth–bro*, *smairf –sprurice*, *trilb–treelilt*, *smeej–leefloag* and *frilg–freezled*). In all, the model produced non-suffixed responses with a strength greater than or equal to that of its weakest correct regular response (a criterion that credits the model for all regular outputs while not saddling it with the dozens of non-suffixed responses weaker than the regular one) an average of 0.9 times per item, 1.4 times per item and 1.3 times per item for the three pseudo-regular classes. In comparison, the average number of non-suffixed responses produced by the subjects in Experiment 3 – a figure that penalises the subjects compared to the model, because they in fact uniformly rated these non-suffixed forms as *worse* than the regular – amounted to 0.2, 0.3 and 0.4 responses per item, respectively.

One might worry whether the difficulties the model of Rumelhart and McClelland had with past-tense forms of novel words was due to some low-level inability to produce unusual sound sequences like those in *ploamphed*. The output module of the model uses a complicated iterative competition to find the set of Wickelphones that best match the output Wickelfeature vector, and then to find the actual phonological string that best matches the Wickelphones, and it is conceivable that some poorly understood aspects of this process militated against converting output patterns for unusual items into cohesive output forms. To test for this possibility, we tested the ability of the model to output the various past-tense forms, independent of its ability to generalise patterns of the stem–past mapping, by simply training it on the novel past-tense forms directly in a past–past identity map. That is, we trained it on suffixed versions of our 60 novel stems, each presented both as input and as target output. We found the model to be capable of outputting 59 of the 60 suffixed forms. For *ploanth*, in the distant pseudo-regular class, it inexplicably produced the suffix in the phonological form [d] rather than the appropriate [t], perhaps because of some quirk in the binding network. Since we have been giving the model credit for incorrect pronunciations of the regular suffix, this deficit is immaterial, and we can be assured that the model is capable of computing output forms for suffixed versions of novel-sounding forms.

Simulation 2

It is possible that the model's difficulties in generalising the regular suffix to the novel-sounding items were due to properties of the novel stems themselves, rather than to problems in applying the regular mapping to novel stems.[6] Indeed, for our human subjects in Experiments 1 and 2, any distaste they had for regularly suffixed novel forms was attributable to properties of the stems themselves. To see if the model, too, simply had trouble with novel stems independent of affixation, with perhaps no increment of difficulty added by the past-tense mapping, we ran a second simulation, using the same training and generalisation set, but training it on a stem–stem identity map (where the input form itself is used as the target output), instead of the stem–past inflectional map.

The model was once again trained for 200 epochs; the response strength for the last 10 epochs exceeded 0.97. When tested on the novel items, the model generalised the identity map equally well to all three verb classes tested, producing 9, 9 and 10 out of 10 identity mappings for the prototypical, intermediate and distant items, respectively. Thus, unlike human subjects, who balk at unfamiliar-sounding stems but experience no further decrement in acceptability when inflecting them, Rumelhart and McClelland's model has no inherent problem treating unusual sounds themselves, but has difficulty in computing novel affixed versions of them.

The model's discrepant performance on identity mapping and past-tense mapping can be attributed to the perfect consistency of the identity mapping, which contrasts with the mutually competing mappings present in the past-tense training set (i.e. a particular input stem pattern might be trained with an identical stem pattern and an affix for a regular verb, but a stem pattern with a different vowel and no affix for an irregular verb).

[6]Note that for models that represent knowledge of the past tense as a mapping between the stem and past forms (e.g. Rumelhart & McClelland, 1986), learning to reproduce the phonology of the stem is part of the task of learning the past-tense form of a verb. For a model of this type, errors in the mapping of the stem and inflections are equivalent.

The past-tense simulation did indeed contain a large percentage of irregular forms, which could interfere with the regular mapping. It is possible, then, that the results of Simulation 1 are not directly comparable with the responses given by our subjects in Experiment 3, because the percentage of irregular verb types is higher than the percentage of irregular verb types in the subjects' vocabularies. To address this limitation, the model was trained on regular verbs only in Simulation 3.

Simulation 3

This time the model was trained on *only* the regular verbs in the corpus used in Simulation 1, eliminating any interference from irregular mappings and thus presenting the most favourable conditions for generalisation of the regular inflection. The model was once again trained for 200 epochs; the response strength for the last 10 epochs exceeded 0.96. The network generalised the regular suffixed past-tense form to 7, 6 and 2 of the 10 items each in the prototypical, intermediate and distant items in pseudo-regular classes.[7]

Thus we see that even in the absence of competition with irregulars, the model of Rumelhart and McClelland has great difficulty generalising regular suffixation to untrained verbs with unfamiliar sound patterns. At first this appears surprising. Pattern associators are affected only by the consistency of the mapping between input and output, and cannot be affected by the labels on the nodes (Fodor & Pylyshyn, 1988; Pinker & Prince, 1988). Thus a competition-free mapping between stem and stem should not be substantially easier than a competition-free mapping between stem and past; the fact that the latter involves nodes labelled with Wickelfeatures that correspond to word-final *-ed* should not have made such a dramatic difference, and one might have expected the model to have shown graded difficulties with more novel forms for both mappings. Two factors may be responsible for the difference: One is specific to the representational scheme used in the model, and the other is due to the nature of the mappings.

First, the output Wickelfeatures representing the end of the stem are word-final for the identity map but word-internal in the past-tense map

[7]Literally the data are 7, 5 and 1, respectively; the totals cited in the text include one form each in the intermediate and distant classes with an incorrect phonological variant of the regular suffix (which should be *ed* after *t* or *d*, *t* after other voiceless consonants, and *d* elsewhere). To be maximally fair to the model, we are counting these errors as correct, because conventional linguistic theories of humans' ability to inflect regular verbs isolate the process of forming the right pronunciation of the *-ed* suffix in a separate, post-lexical phonological component (see Pinker & Prince, 1988).

(because a suffix intervenes between the final segments and the end of the word). In Rumelhart and McClelland's coding scheme, the position of a phonological feature within a word is not represented explicitly (only its position relative to the preceding and following feature), which poses problems in reassembling the word from its activated features and in efficiently learning that the form of the English past tense depends more heavily on specific features at the ends of words than features at the beginnings or middles. Rumelhart and McClelland dealt with these problems by coding word boundaries as if they were phonemes, but with one important difference, noted by Lachter and Bever (1988). In their scheme, real phonemes in particular positions like *d* do not directly define any Wickelfeatures; the Wickelfeatures that encode them correspond only to the features that they are composed of. In contrast, the end-of-word pseudophoneme defines a unique feature itself, which combines with all logically possible pairs of features preceding it into 100 different Wickelfeatures. Thus a sequence like *m-f*-word boundary is represented by 16 Wickelfeatures, any of which is uniquely sufficient to indicate that the feature sequences contained in *m-f* occur at the end of a word. In contrast, for the sequence *m-f-t*, no single Wickelfeature encodes the information that any of *m-f*'s feature sequences precedes *t*; each sequence is represented as preceding one of *t*'s features. Indeed, an additional aspect of the coding scheme introduces more uncertainty because only a subset of the *m-f-t* feature sequences are represented as unique Wickelfeatures, namely those involving features of *m* defined over the same dimensions as features of *t* (thus the fact that *m-f-t* contains an interrupted-*f*-interrupted sequence represented by turning on a unique Wickelfeature, but the fact that it contains a nasal-*f*-unvoiced sequence is not; in contrast, both interrupted-*f* and nasal-*f* are explicitly represented as preceding a word boundary, each by its own Wickelfeature). Given training at a level of less than maximum strength on all links between input Wickelfeatures and their output counterparts, there could be more redundancy, and hence more reliable reconstruction of the string, in novel verbs presented to the net trained to do the identity map than in novel verbs presented to the net trained to do the regular past-tense map.

Secondly, the purely regular past-tense mapping actually does involve some competition in Rumelhart and McClelland's model because the suffix assumes three forms (*t*, *d* and *ed*) and the model is charged with computing the correct one, not just any version of an alveolar stop. Selection of the proper variant cannot be accomplished by any single link between a Wickelfeature and the Wickelfeatures corresponding to a given pronunciation of the output affix, but depends on summing the effects of combinations of such links: Wickelfeatures representing an unvoiced consonant at the end of a stem should trigger *t* in the output unless there are other Wickelfeatures representing an alveolar stop at the end of a stem, in which case the word-final *t* feature sequences in the output are suppressed and *ed* ones are activated. Novel forms like *ploamph* may contain Wickelfeature combinations that have not been encountered before and hence that have not been taken into account in setting the link weights to ensure cohesive output forms for all inputs. (Although we counted incorrect versions of the regular suffix as correct, the many-to-many mapping between input and output for the regular map could interfere with producing any version at all for novel items.) In contrast, the identity map is perfectly consistent – each input Wickelfeature turns on its twin in the output, and novel combinations of Wickelfeatures can be responded to properly, as long as each component Wickelfeature link (or a redundant enough subset to allow reconstruction of the word) has been trained on previous inputs. In perfectly consistent mappings such as the identity mapping, the network does not actually have to learn to map patterns onto patterns but can map features onto features. In a sense, the representations used in such a mapping are not distributed but local.

Discussion

The simulations demonstrate that a prominent associative model capable of occasionally generalising irregular patterns to novel similar items can also generalise the regular pattern to novel similar items, but has great difficulty in simultaneously generalising the regular pattern to novel dissimilar items. This was even true for the simulation in which the training set consisted *only* of regular verbs. It is not attributable to any problem in handling novel input or output stem patterns themselves. The lack of a difference between the generalisation patterns of simulations with and without competition from irregulars suggests that the model's generalisation difficulty is not just due to competition with a specific irregular pattern, but also to the sheer dissimilarity of the novel items to the trained items when faced with a mapping of even moderate complexity.

Because Rumelhart and McClelland's model has far less linguistic experience than any English-speaking adult, the comparisons we have been making might not be fair to the model. Perhaps one should examine what the model's performance would be if it were trained on a large number of English regular verbs, comparable to the number that adults have learned. Such a set might exemplify broader regions of phonological space, allowing the model to improve its performance on peripheral verbs. There are two reasons to doubt that its behaviour would duplicate that of our adult subjects, however. First, because distant pseudo-regular items like *ploamph* are phonologically illicit in English, increasing the number of real English vocabulary items would not cover these regions of phonological space. Secondly, because the model was below ceiling even for prototypical pseudo-regulars, it is possible that a larger training set would improve performance for these items, leaving the model's steep generalisation gradient intact.

An alternative would be to compare the model's performance with the current training set to the behaviour of children, who have been exposed to a similar number of regular verbs (see Marcus et al., 1992). Though we lack data comparable to those of Experiment 3 for children, Marcus et al. (1992) note that in several regards even children appear to apply the regular suffix to novel stems without requiring them to have high degrees of similarity to frequently encountered stems in parental speech. Marcus et al. analysed over-regularisation errors like *comed* in the spontaneous speech of 19 children. They found that the over-regularisation rates for different irregular verbs did not correlate with the number and frequency of similar regular verbs in their parents' speech (that is, frequent exposure to forms like *blinked* did not make corresponding over-regularisation errors like *drinked* more common). Moreover, the child with the richest data set, Abe, over-regularised 90% of the irregular verbs in his vocabulary at least once, together with many of his own unusual-sounding inventions such as *eat lunched, bonked, borned, axed, fisted* and *poonked.* These phenomena suggest that children, like adults, might possess a regular inflectional process that has wide applicability across verb sounds.

GENERAL DISCUSSION

The results of the experiments and simulations and the conclusions they suggest can be summed up as follows. Replicating Bybee and Moder (1983), we find that people's willingness to generalise inflectional patterns from existing irregular verbs to novel ones containing a relevant vowel falls off with the global similarity of the new verb to the existing verbs showing that pattern. In contrast, the regular past-tense suffix does not appear to depend for its generalisation on similarity to existing regular verbs. In absolute terms, regularly suffixed past-tense forms were consistently given high average ratings (4.8 or greater on the 7-point scale) and were produced a majority of the time, even for classes of verbs that were most strongly pulled to irregular generalisations, and even for classes of verbs that resemble no English regular verb. In relative terms, regularly suffixed past forms showed small differences in ratings and production probabilities depending on similarity to existing English regular verbs, but these differences were frequently inconsistent in direction and not statistically significant, and when they did occur they were attributable to independently measured perceptions of the phonological difficulty of the sounds of the stems themselves. Thus although the experiments do not rule out the possibility that there are associative-memory effects on generalisations of regular morphology, they provide no evidence for such effects, and if they exist at all, their effect on people's linguistic behaviour appears to be small.[8]

The subjects' behaviour in generalising irregular patterns to verbs with different degrees of resemblance to irregulars is fairly consistent with Rumelhart and McClelland's connectionist model, in that both the subjects and the model easily handled the phonological representations of the stems themselves, and showed occasional generalisation of the inflectional patterns, dependent on similarity. However, the model differed from the subjects in its treatment of generalisations of regular patterns to verbs with different degrees of resemblance to regulars: Whereas the model had no trouble with unusual stems themselves but had great trouble inflecting them with the regular suffix, humans were sensitive to the unusualness of the stems themselves but produced regularly suffixed versions of them reliably, and their suffixation process did not appear to degrade the quality of the suffixed form relative to the stem.

In the rest of the paper, we evaluate current theories of the psychology of inflection in terms of their ability to account for the data presented in these experiments and related phenomena. As discussed in the Introduction, the generalisation of irregular morphological patterns refutes the theory that irregular forms are stored in a rote list, but are not capturable by rules with necessary and sufficient conditions. Thus we suggest that irregulars are stored and generalised in a memory system containing a level of representation in which the phonological properties of words are superimposed, fostering graded generalisation by similarity. Two kinds of models will now be considered: pure connectionist models like that of Rumelhart and McClelland (1986) and Plunkett and Marchman (1990), consisting of a single associative memory system, and the modified traditional model discussed by Pinker and Prince (1988; 1991), Pinker (1991), Marcus et al. (1992) and Kim et al. (1991b), in which regular generalisations are not similarity-based analogies, but are computed by a rule in the sense of a concatenation operation.[9]

---

[8]Much of the uncertainty as to whether there exists any effect of similarity on regular generalisations owes to the design of these experiments, where to ensure that intermediate and distant items did not resemble any of the thousands of English regular verbs, we were forced to test highly unusual and illicit sound patterns, which appeared to introduce their own effects. Replications of this experiment with materials from some other language that allows the creation of seldom-regularised but phonologically - licit stem patterns might resolve the uncertainty.

[9]This hybrid model is related to earlier proposals in the literature of generative linguistics that true morphological rules are to be distinguished from "redundancy rules", which capture varying degrees of systematicity in lexical memory without freely licensing productive extensions to new forms. See Jackendoff (1975), Aronoff (1976) and Lieber (1980) for examples.

Relevance to Newer Connectionist Models

We have seen that Rumelhart and McClelland's model accounts reasonably well for the semi-productivity of irregular patterns, based on similarity to trained exemplars, but has problems generalising the regular pattern to novel forms which are very different from existing forms. McClelland (1988), responding to the first discussions of such problems in Pinker and Prince (1988), speculated that they are due to technical limitations of Rumelhart and McClelland's model, primarily its lack of a hidden layer of nodes between input and output. Plunkett and Marchman (1991) also attribute the problems of the model to its early two-layer architecture. In this section, we evaluate these conjectures by examining three new kinds of connectionist models, each augmenting Rumelhart and McClelland's model in a different way.

Single Networks with One Hidden Layer. Plunkett and Marchman (1990, 1991) report two sets of simulations of a single-hidden-layer network trained on a set of English-like artificial stems. Neither report supplies data relevant to the models' ability to generalise to novel-sounding stems, however. Plunkett and Marchman (1991) did not test the model's generalisation to novel stems at all, and Plunkett and Marchman (1990) used only stems that were similar to the ones the models had been trained on. Specifically, the model's input feature vector implicitly defined a space of about 250,000 possible combinations of features into stems; constraints on possible consonants, vowels and CV sequences reduced this space to 14,400. Plunkett and Marchman assembled a dictionary of 1000 verbs from this space, trained the model on a regular suffixation mapping using 458 of them, and tested its generalisation on 50 novel stems from the same dictionary. Thus the training and generalisation items were drawn from the same small region of the space of possible forms and hence were similar to one another, analogous only to our prototypical pseudo-regular items. The model's successful generalisation to these new items (90% for large vocabulary sizes) does not tell us how well it would generalise to items selected so as to be peripheral.

Egedi and Sproat (1991; see also Sproat, 1992) devised a connectionist model that had a similar overall structure to Rumelhart and McClelland's, trained and tested it on a comparable set of English verbs, and augmented it with a hidden layer whose weights were adjusted by the standard error back-propagation algorithm, as advocated by McClelland and by Plunkett and Marchman. The model also enjoyed a more realistic phonological representation (not Wickelfeatures), and a mechanism converting the output node activations to pronounceable words that was more powerful than Rumelhart and McClelland's binding networks. Like Rumelhart and

McClelland's model, it was able to capitalise on the similarity-based generalisation afforded by its distributed representation to guess the past-tense forms of several novel irregular verbs, like *stick, creep, sweep* and *ring*. However, also like Rumelhart and McClelland's model, it generalised poorly to novel English regular verbs. It produced correct suffixed forms for 62% of the new regular forms it was tested on, left 6% unchanged, assimilated 6% to other verbs in the training set (e.g. *train–trailed, spoke –smoked, glow–glanced*) and rendered bizarre distortions for 25% (e.g. *conflict–conflafted, wink–wok, yield-rilt, satisfy–sedderded, quiver–quess*).

Similarly, Seidenberg and McClelland's (1989) three-layer model of word recognition showed similarity effects among existing words that allowed it to account for an impressive amount of human response time and error data. However, it suffered from significant problems in outputting correct pronunciations of novel items, suggesting that it had inherent difficulties in generalising its knowledge of spelling–sound correspondences to words that are unlike those in its training history (Besner, Twilley, McCann, & Seergobin, 1990; Coltheart, 1991).

The generalisation limitations of the Egedi–Sproat and Seidenberg–McClelland models suggest that the problems the Rumelhart–McClelland model had in generalising its knowledge of the regular inflection is not due solely to its two-layer architecture.. A likely diagnosis is that any PDP model will have difficulty with the requirements of natural language morphology if it uses a single associative network to compute both regular and irregular mappings. The problem is that a single mechanism, undivided into modules specialised for the very different kinds of computation each involves, is likely to end up as a compromise between the conflicting demands of the two mappings, handling neither one properly. The regular inflection requires a coarse similarity metric such that any new item can be treated as being similar to previously known items, whereas irregular inflections require a finer-grained similarity metric, allowing similar items to have different mappings. A key property of nets related to this issue is the *sparseness vs denseness* of the representation.

Automatic generalisation occurs in PDP networks as a direct consequence of two items sharing representational units. If a novel item shares many representational units with a known item (i.e. is similar to a known item), the mapping that has been learned for the known item is likely to generalise to the novel item. The likelihood of generalisation of a particular mapping to a novel item thus increases with the similarity of the novel item to other items with the given mapping and decreases with the novel item's similarity to items with different mappings, all else being equal. For example, a network that has learned the past-tense forms of English verbs would likely inflect a novel item such as *spling* as *splung* because it is likely to share many representational units with items like *string, swing, cling* and *spin* and not many units with items that have different mappings. A novel item such as *clow*, however, would be less likely to elicit a past-tense form such as *clew* even though it is likely to share many representational units with items such as *grow, blow, know* and *throw*, because it would also share many units with items that have a different mapping such as *tow, row, glow* and *flow*. In networks with hidden layers, not only the overlap among units in the input representation, but also the overlap among units in the hidden layer, is relevant to the networks' ability to learn different mappings for similar items. The amount of overlap between two items at the input level is completely determined once one picks a representational scheme, but the amount of overlap in the hidden layers depends on the training conditions, architecture and learning rule.

The sparsest kind of representation is a localist one. Localist representations are orthogonal and thus do not lead to any generalisation from one item to another. The possibility of sparse representation of a set of items increases as one increases the dimensionality of the space. The dimensionality of a space, however, does not uniquely determine how sparse the representations are – the mapping from objects to vectors is also relevant. For example, in an eight-dimensional space, rather than representing each of eight objects on its own node, a maximally sparse coding, one could choose a base-two representational scheme in which the eight items are represented on three units, a denser coding showing considerable overlap. Across such representations, the ability to learn different mappings for similar items and the ability to generalise a given mapping to a similar item are inversely related. Relatively sparse representations allow there to be distinct mappings for similar items, but lead to poor generalisation. Sloman and Rumelhart (1991) and French (1991) have shown how the use of relatively sparse representations leads to a reduction in interference between different mappings. Denser representational schemes lead to better generalisation but have more difficulty learning different mappings for similar items, because of increased interference from other mappings.

Computing the regular suffix, the irregular changes, and the preserved portion of the stem each require different degrees of sparseness in their optimal input representation. Consider a completely regular inflectional mapping first, like the regular-only simulation (a real-life example is the English *-ing* inflection, which shows no irregularity). To compute the presence of the affix, a maximally dense representation is ideal – indeed, if all the input nodes funnelled into a single hidden node, which connected to output nodes representing the affix, generalisation of the regular suffix to new forms would be perfect. However, a proper past-tense form needs the stem to be copied, and does not just consist of a bare suffix, so it benefits from a representation that is sparse at the level of phonemes (one unit per phoneme in a position) but dense at the level of the stems, so that, for example, *ploamph* can be produced based on prior training of *plant, coast* and *laugh*. Finally, the irregular mappings, because they can differ in virtually any segment from one another (e.g. *know–knew/go–went; tear–tore/tell–told; spring–sprang/string–strung/bring–brought*) would be happiest with a very sparse representation. A representation where every stem was represented independently of the others, as in the traditional rote-list theory, would lead to no errors; introducing some degree of overlap would give it the ability to make the irregular generalisations that are occasionally called for. (In Appendix 2, we discuss reasons why languages tend to have irregular forms to begin with.)

Modular Systems. The differing computational demands of the various sub-problems in inflectional marking, when set against the properties of distributed representation, suggest that architectures that use separate networks for mapping the regular and irregular inflection are likely to be more successful than those that try to achieve both types of mappings within one network such as Rumelhart and McClelland's model and multilayer augmentations of it. Jacobs, Jordan, and Barto (1991; see also Rueckl, Cave, & Kosslyn, 1989) have shown that temporal and spatial cross-talk (interference between mappings trained at different times, or for different portions of the space of outputs, respectively) can be reduced through the use of *modular* connectionist architectures. In such systems, inputs are shunted to one of several distinct networks, possibly with innately different architectures, depending on properties of the input that signal the nature of the computation that the items should undergo. Jacobs et al. showed that modular architectures lead to better learning than single networks when learning two tasks with distinct computational demands. This result is reminiscent of more standard models of grammar, including the modified traditional view laid out in Pinker and Prince (1991), Pinker (1991) and Kim et al. (1991b), where storage of the lexical entries, computation of regular inflection, and linkage of irregular forms to their stems are computed by different modules of the system, unimpaired by cross-talk among them. Whether a modular system of the kind that Jacobs et al. describe could in fact learn an optimal partitioning of the past-tense mapping among networks is an interesting open question. An important issue is whether their mechanism for gating inputs to different modules, a kind of pattern associator, might have difficulty knowing what to do with types of items it had never been trained on, introducing generalisation failures at the gating level even if a module capable of generalising the pattern existed within the system.

An Architecture Customised for the English Past-tense Mapping. The empirical and computational demands for modular architectures for past-tense forms is also shown indirectly in MacWhinney and Leinbach's (1991) past-tense network. MacWhinney and Leinbach's goal was to show that there existed a connectionist model that avoided the problems of Rumelhart and McClelland's model pointed out by Pinker and Prince (1988). MacWhinney and Leinbach's model successfully learned the past-tense forms of the verbs in the training corpus but was not tested on its ability to generalise the regular inflection; it seems likely, however, that the model would generalise the regular inflection quite well.

The model had two hidden layers (rather than zero as in Rumelhart and McClelland's model, or one in Plunkett and Marchman's simulations), each with 200 hidden nodes. Note that because the model was trained on only 118 irregular verbs, there are enough nodes to represent each lexical item locally inside the network, in its own grandmother cell. Furthermore, each verb is represented twice in the input in non-overlapping sub-vectors, one consisting of left-justified representation of the full ordered phonological string, the other consisting of a right-justified representation of the rhyme (final vowel and consonants). Since the sub-regularities in the irregular clusters are defined mainly over their rhymes (Pinker & Prince, 1988), the right-justified representation is ideal for picking up the sub-regularities supporting similarity-based generalisation, while the left-justified full representation is then available to code idiosyncratic lexical differences (e.g. between *string, spring, bring*, and regular *blink*). In a sense, the right-justified representation serves as a surrogate for the phonological representation of the stem, entering into similarity-based generalisations, and the left-justified representation serves as a surrogate for the traditional notion of a verb's lexical entry as the stored locus of its idiosyncratic information.

Most tellingly, the model also has an identity map hardwired into it, in the form of a separate pathway connecting each input node with its twin in the output with a maximally weighted innate link. This makes the task of learning the mapping between the stem and its preserved portion in the past-tense form much easier: MacWhinney (pers. comm., 1991) notes that before they implemented this copying network, the model (like those of Rumelhart and McClelland and Egedi and Sproat) was prone to producing distortions of the stem rather than exact copies, and that when it had finally been trained to avoid such "garbage", it had devoted so many units to the sound-to-sound mapping for stem-copying that there were not enough to learn the change in the past-tense form. Since regular verbs require that 100% of the stem be copied, the copying network can be regarded as an innate component adapted to the requirements of the regular past-tense mapping (see Pinker and Prince, 1988 p. 168, who explicitly suggest that this architecture is a straightforward implementation of an identity rule), and, to a lesser extent, to the aspect of the irregular mapping that traditionally has been ascribed to rules (see Pinker & Prince, 1991; Pinker, 1991), namely the preservation of much of a stem in its irregular past-tense form. The hidden-layer associator, then, need only learn the idiosyncrasies of the irregular mappings (and the three phonological variants of the regular suffix, which the model treats as a kind of sub-regularity).

In sum, MacWhinney and Leinbach appear to have reinvented, through trial and error, counterparts to the modular distinctions required by traditional linguistic theory when they augmented Rumelhart and McClelland's model with new kinds of innate nodes and connections. Very roughly, left-justified features appear to be designed to do the work of lexical entries; right-justified features, phonological representations; identity bypass route, regular rule and the rule-like components of the irregular mappings. (In addition, they added semantic features to the input vector, a crude way of approximating the effects on inflection of a word's morphological structure: see Kim et al., 1991a, 1991b, and below.) Each of these tactics helped their model avoid one of the problems noted by Pinker and Prince. As a result, they ended up with a model with an architecture and representation that is custom-designed for the English past-tense mapping, not a generic connectionist problem-solver. In the next section, we will show that even these modifications appear to be psychologically unrealistic stop-gaps when the larger context of morphological representation and processing is examined.

Relevance to Rule-based Models

The data from these experiments and simulations suggest that humans generalise irregular patterns only to items highly similar to existing items, but can generalise the regular pattern accurately even to highly dissimilar items, an ability that contrasts with homogeneous connectionist models. This pattern is consistent, however, with a modified traditional model (Pinker, 1991; Pinker & Prince, 1991) distinguishing three data structures: a lexicon that stores the idiosyncratic bundle of information that defines a word, including irregular forms; a representation of the phonological content of stem–past pairs, in which similar pairs partially overlap, fostering analogical generalisation by similarity; and a regular rule, that concatenates an affix to any word bearing the "verb" symbol, ignoring its phonological content. As mentioned in the Introduction, it remains an open question in this theory whether regularly inflected items are ever redundantly stored in the lexicon (see Ullman & Pinker, 1991, for discussion and data). Crucially, generalisations of the regular pattern do not *depend* on such storage within a rule-using theory, though they are compatible with such storage. In such a model, all three modules may or may not be modelled by some kind of connectionist network, though only for the phonological associative memory would the special generalisation properties of standard distributed connectionist networks be exploited (the others, if modelled by connectionist mechanisms at all, would likely be implementations of traditional symbol manipulation in more or less localist networks).

There are reasons independent of successful generalisation to novel items that this modular model appears to be a more psychologically realistic model of inflection. A general advantage of modular systems is that the components can interact with other processes in distinct ways (Jacobs et al., 1991). The three-way distinction between lexical storage, associative generalisation of phonological patterns and regular suffixation was not explicitly proposed to handle the generalisation patterns that we document in these experiments; it is independently motivated by qualitative input–output requirements of the full grammatical system, and the generalisation patterns are a direct consequence.

First, non-associative, quasi-rote storage of lexical items is not only helpful in learning idiosyncratic past-tense mappings (e.g. *go–went*, and the difference between *bring–brought* and *spring–sprang*); it is necessary for homophones with different past-tense forms [e.g. *lie–lay* (recline) *vs lie–lied* (fib); *meet–met vs mete–meted; wring–wrung vs ring–rang*; see Kim et al., 1991b). Although MacWhinney and Leinbach (1991) propose to distinguish such pairs using semantic features, this falsely predicts that irregular generalisations will tend to be more likely for items sharing a semantic feature with an existing irregular, just as irregular generalisations are more likely for items sharing a phonological feature. That is, words with similar meanings are predicted to be likely to have similar past-tense mappings, for the same reason that words with similar sounds have similar past-tense mappings. However, the English system shows no hint of such generalisations either in its history or in modern speakers (see Kim et al., 1991a, 1991b).

More generally, the lexicon must represent the relationship between phonological forms and their semantic representations: Given a form such as *walk* or *sing*, people can retrieve the form's meaning, and vice-versa. Unlike the past-tense system which has a few arbitrary mappings and some idiosyncratic sub-regularities but is largely regular, the mapping between phonological form and meaning is almost completely arbitrary (with minor exceptions such as are involved in phonetic symbolism and phonesthemes). Similarly, the relationship between the form class of a word and its phonological form is almost entirely arbitrary. Distributed representations alone are not suited to representing lexical information, because a lexical entry consists of connections between phonological, semantic, morphological, and syntactic information and the mappings that hold between these types of information are mostly arbitrary, especially between phonological information and the other types of information. This state of affairs is suited to a localist or symbolic representation and ill-suited to distributed representations which make use of analogical processes.[10] Crucially, given the need for such representations for basic form–meaning pairings, idiosyncrasies in past-tense forms are an automatic possible consequence: If you can arbitrarily pair the sound /fid/ with the lexical entry corresponding to "serve food to", you can just as easily pair the sound /fɛd/ with the entry corresponding to "the past tense of serving food to" (see Appendix 2 for a discussion of why languages might occasionally exercise this option).

Independent qualitative evidence that irregulars are stored in memory (despite their partial generalisability) is that they behave like stored lexical items, and unlike otherwise similar regular forms, in their interaction with processes of derivational word formation. For example, in English one can form a lexical compound by joining two words, as in *mud–eater*. Kiparsky (1982a; 1982b) notes that irregularly inflected items can easily appear within compounds, but regularly inflected items cannot: compare *mice–infested* to **rats–infested*, *purple–people–eater* to **purple–persons–eater*, *men–bashing* to **guys–bashing*. Senghas, Kim, Collins, and Pinker (1991), using data from subjects who rated a large number of novel compounds, confirmed that irregularity has a robust and unconfounded positive effect on the naturalness of compounds containing plurals. Moreover, similar effects appear with the fully regular plural inflection *-s* in Dutch (Senghas et al., 1991) and in German (Clahsen, Rothweiler, Woest, & Marcus, in press). Gordon (1985; 1986) has shown that even very young children respect the regular–irregular difference when forming new compounds in response to experimental elicitation: they call "a monster who eats mice" a *mice-eater*, but call "a monster who eats rats" a *rat-eater* not a *rats-eater*. This difference between regular and irregular forms is explainable if irregularly inflected items behave just like any other word stored in memory, and hence can feed the compounding process, but regularly inflected items are formed downstream in the information flow from lexicon to syntax, too late to enter the lexical compounding process. This account directly predicts that irregular forms are stored in memory, and that whatever generalisation they support should be via analogies to similar memorised items, unlike rule-governed processes which can be completely general.

---

[10]Hinton, McClelland, and Rumelhart (1986, p. 78) point out: "Every representational scheme has its good and bad points. Distributed representations are no exception. Some desirable properties arise very naturally from the use of patterns of activity as representations. Other properties, like the ability to temporarily store a large set of arbitrary associations, are much harder to achieve."

As for the regular affixation process, again there is an independent qualitative earmark distinguishing it from memory and irregular patterns. A major grammatical difference between regular and irregular inflection is that regular inflection is a *default* process and applies to all forms unless explicitly blocked by an irregular root. This includes both novel verbs like *faxed*, and any form not based on a verb root like *spitted/*spat the pig, ringed/*rang the city* and *grandstanded/*grandstood*. Such forms *cannot* receive irregular inflection even though they are maximally close (i.e. identical) phonologically to irregulars (Kim et al., 1991b; Kiparsky, 1982a; 1982b; Mencken, 1936; Pinker & Prince, 1988), because irregularity is a property of verb roots, not verb sounds. A verb intuitively based on a noun has a noun root, represented in the verb's morphological structure (e.g. $[[\text{grandstand}]_{\text{N}}]_{\text{V}}$), and a noun root cannot be listed in the lexicon as having an irregular past, because nouns cannot have past-tense forms listed for them at all; the regular rule applies to such verbs as the default.

Because Rumelhart and McClelland's model represents only the phonology of stems, it is incapable of discriminating forms like *rang* and *ringed*; this was the first generalisation failure mentioned in the Introduction. MacWhinney and Leinbach's tactic of inserting semantic features into the input vector fails to address the problem outside the confines of their restricted set of verbs for two reasons. One is that Kim et al. documented that morphological structure (noun *vs* verb root), not semantics, is the predictive factor in the regularisation of such verbs when the two are unconfounded. But even if the relevant morphological structure was encoded in the input vector of a pattern associator, say, with a unit that was turned on for verbs with noun roots, the full set of facts would not be captured parsimoniously. Within the noun system, related principles apply, and the regular suffix *-s* is applied in a variety of circumstances in which irregular forms do not exist or cannot apply for grammatical reasons (see Pinker & Prince, 1991). These include not only neologisms like *modems*, but nouns from names (*The Manns/*Menn are coming over for dinner; I'm sick of all the Mickey-Mouses/*Mickey-Mice in this administration*), nouns lacking heads (*Walkmans, low-lifes, tenderfoots*), nouns derived from verbs (*With a few quick wolfs/*wolves, he devoured his lunch*), and nouns derived from phrases (*While checking for sexist writing I found three "every man"s/*"every men" on the first page*). The key point is that a regular rule need not be stipulated to apply to each and every one of these circumstances (many quite rare and unusual); rather, it is the default process that speakers can use *whenever* an irregular is not applicable, for any reason. And, crucially, given a process designed to be available across-the-board, not to any predefined set of items, the ability of the process to apply freely to forms with unusual sounds like *ploamph* comes out for free, and does not have to be approximated by tweaking or overtraining a network whose architecture is inherently suited to associating patterns with patterns.

Modern English presents us with the unfortunate methodological confound that the inflectional process that behaves in a rule-like way also applies to a very large number of verbs. This has contributed to the partial successes of the single-net models that have been trained on the past-tense mapping, because the large number of regularly inflected verbs that exist in English can lead to a state of training in which any new verb is similar enough to some known verbs to undergo the dominant mapping. But cross-linguistic and historical comparisons suggest that the two properties can be unconfounded, and that when they are, it is the default character of regular inflection that can lead to a majority of verbs being regular, not vice versa. For example, the two properties are unconfounded in Modern Arabic, which has two modes of plural formation (see McCarthy & Prince, 1990). The "broken plural" is the most common set of forms, but is limited to various classes of similar canonically shaped nouns. The "sound plural", in contrast, behaves like a default in that it applies indiscriminately to all non-canonical forms, such as proper names, transparently derived nouns such as deverbals and diminutives, non-canonical or unassimilated borrowings, and the names of letters of the alphabet, which are mostly non-canonical; all these forms, though, constitute a minority of Arabic nouns. Because the items which the default inflection applies to are relatively few and can come from many areas of phonological space, it is likely that a novel item such as a name or borrowing will not be very similar to known items and thus not be properly inflected by an analogy-based mechanism, but the regular sound plural can apply nonetheless. Similarly, Clahsen et al. (1992) and Marcus (1992) show that the German *-s* plural, though applying to a tiny minority of nouns, acts like a fully regular default plural by a variety of linguistic and psychological tests.

Indeed, the ancestors of Modern English might have had this property. The strong verbs in Modern English are fossils of Indo-European ablaut classes, which used to be the form of past-tense inflection applied to most verbs. The weak *-ed* suffix was introduced in Proto-Germanic and was used there and in its immediate descendants for derived forms and borrowings (Pyles & Algeo, 1982). Two peculiarities of Modern English are that it borrowed many of its verbs from French and Latin (perhaps 60%), and that it formed large numbers of its other verbs from nouns (perhaps another 20%) – both of which are default circumstances that call for the weak suffix.[11] Thus the preponderance and heterogeneity of regular verbs

[11]These rough estimates, provided by Michelle Hollander, are based on a sample of 200 verbs drawn randomly from the approximately 4000 in Francis and Kučera (1982).

in English might be an epiphenomenon of the rule-like default nature of the weak suffixing process, rather than vice versa. If so, it would be getting it backwards to stimulate the mental process of regular inflection by exploiting the statistics of Modern English, and the strategy should fail outright in modelling regular processes like the Arabic sound plural, the German *-s* plural, and the ancestor of the weak past-tense suffix of Modern English, where the most frequent inflection does not correspond to the default inflection.

In sum, the current study found that people's likelihood of producing – and judgements in accepting – irregularly inflected forms varied strongly as a function of a stem's phonological distance from existing irregular clusters. However, regularly inflected forms were produced and accepted at high levels for all kinds of verbs, and did not vary reliably as a function of the stem's distance from "regular" clusters. Any model of past-tense inflection must account for the different manners in which irregular and the regular inflection are generalised. Models motivated both by the full set of evidence on inflection (including grammatical constraints in English, cross-linguistic comparisons and historical development) can account for the difference using independently motivated distinctions between lexical memory and rule, if the memory component is enriched by attributing to it certain analogy-fostering properties reminiscent of distributed connectionist models. Unitary, single-net connectionist models do not naturally show this difference, other than by *ad hoc* customisation that isolates the model from the larger set of data on the psychology of inflection. Thus modular models, combining a memory module with both local and distributed representations, and a module carrying out traditional rule-like processes, appear to be the most promising architecture for inflection, and might point towards resolution of some of the debates between connectionist and symbolic generative-grammar theories of the psychology of language.

Manuscript received October 1991
Revised manuscript received March 1992

REFERENCES

Aronoff, M. (1976). *Word formation in generative grammar*. Cambridge, Mass.: MIT Press.

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

Besner, D., Twilley, L., McCann, R.S., & Seergobin, K. (1990). On the association between connectionism and data: Are a few words necessary? *Psychological Review, 97*, 432–446.

Borowsky, T. (1989). Structure preservation and the syllable coda in English. *Natural Language and Linguistic Theory, 7*, 145–166.

Brown, R. (1973). *A first language: The early stages*. Cambridge, Mass.: Harvard University Press.

Bybee, J.L. (1985). *Morphology: A study of the relation between meaning and form*. Philadelphia, Penn.: Benjamins.

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

Bybee, J.L. & Slobin, D.I. (1982). Rules and schemes in the development and use of the English past tense. *Language, 58*, 265–289.

Cherry, C. (1966). *On human communication*. Cambridge, Mass.: MIT Press.

Chomsky, N. & Halle, M. (1968). *The sound pattern of English*. New York: Harper and Row.

Clahsen, H., Rothweiler, M., Woest, A., & Marcus, G.F. (1992). Regular versus irregular inflection in the acquisition of German noun plurals. *Cognition, 45*.

Coltheart, M. (1991). Connectionist modeling of human language processing: The case of reading. *PSYCHOLOQUY: Refereed Electronic Journal of Peer Discussion, 2*, 9.3.1.

Egedi, D.M. & Sproat, R.W. (1991). *Connectionist networks and natural language morphology*. Unpublished manuscript, Linguistics Research Department, AT & T Bell Laboratories, Murray Hill, NJ.

Fodor, J.A. & Pylyshyn, Z. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition, 28*, 3–71.

Francis, N. & Kučera, H. (1982). *Frequency analysis of English usage: Lexicon and grammar*. Boston, Mass.: Houghton Mifflin.

French, R.M. (1991). Using semi-distributed representations to overcome catastrophic forgetting in connectionist networks. In *Proceedings of the Thirteenth Annual Conference of the Cognitive Science Society*. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Gordon, P. (1985). Level-ordering in lexical development. *Cognition, 21*, 73–93.

Halle, M. (1973). Prolegomena to a theory of word formation. *Linguistic Inquiry, 4*, 3–16.

Halle, M. & Mohanan, K.P. (1985). Segmental phonology of modern English. *Linguistic Inquiry, 16*, 57–116.

Hinton, G.E., McClelland, J.L., & Rumelhart, D.E. (1986). Distributed representations. In D.E. Rumelhart, J.L. McClelland, & the PDP Research Group, *Parallel distributed processing: Explorations in the microstructure of cognition. Vol 1: Foundations*. Cambridge, Mass.: Bradford Books/MIT Press.

Jackendoff, R.S. (1975). Morphological and semantic regularities in the lexicon. *Language, 51*, 639–671.

Jacobs, R.A., Jordan, M.I., & Barto, A.G. (1991). Task decomposition through competition in a modular connectionist architecture: The what and where vision tasks. *Cognitive Science, 15*.

Kim, J.J., Marcus, G., Hollander, M., & Pinker, S. (1991a). Children's inflection is sensitive to morphological structure. *Papers and Reports in Child Language Development, 30*, 39–46.

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

Kiparsky, P. (1982a). From cyclical to lexical phonology. In H. van der Hulst & N. Smith (Eds), *The structure of phonological representations*. Dordrecht: Foris.

Kiparsky, P. (1982b). Lexical phonology and morphology. In I.S. Yang (Ed.), *Linguistics in the morning calm*, pp. 3–91. Seoul: Hansin.

Kučera, H. & Francis, N. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.

Lachter, J. & Bever, T.G. (1988). The relation between linguistic structure and associative theories of language learning – A constructive critique of some connectionist learning models. *Cognition, 28*, 195–247.

Lederer, R. (1989). *Crazy English*. New York: Pocket Books.
Lieber, R. (1980). *On the organization of the lexicon*. Doctoral dissertation, Department of Linguistics and Philosophy, MIT.
MacWhinney, B. & Leinbach, J. (1991). Implementations are not conceptualizations: Revising the verb learning model. *Cognition, 40*, 121–157.
Mandelbrot, B. (1954). Structure formelle des textes et communication. *Word, 10*, 1–27.
Marcus, G.F., Brinkmann, U., Clahsen, H., Wiese, R., Woest, A., & Pinker, S. (1992) *German inflection: The exception that proves the rule*. Unpublished manuscript, Dept. of Brain and Cognitive Sciences, MIT, Cambridge, Mass.
Marcus, G.F., Pinker, S., Ullman, M., Hollander, M., Rosen, T.J., & Xu, F. (1992). Overregularization in language acquisition. *Monographs of the Society for Research in Child Development, 57* (4, Serial No. 228).
McCarthy, J.J. & Prince, A.S. (1990). Foot and word in prosodic morphology: The Arabic broken plural. *Natural Language and Linguistic Theory, 8*, 209–283.
McCarthy, J.J. & Prince, A.S. (in press).*Prosodic Morphology*, Cambridge Mass.: MIT Press.
McClelland, J.L. (1988). Connectionist models and psychological evidence. *Journal of Memory and Language, 27*, 107–123.
Mencken, H. (1936). *The American language*. New York: Knopf.
Pinker, S. (1991). Rules of language. *Science, 253*, 530–535.
Pinker, S. & Prince, A. (1988). On language and connectionism: Analysis of a parallel distributed processing model of language acquisition. *Cognition, 28*, 73–193.
Pinker, S. & Prince, A. (1991). Regular and irregular morphology and the psychological status of rules of grammar. In L. Sutton (Ed.) *Proceedings of the 17th Annual Meeting of the Berkeley Linguistics Society*. Berkeley, Calif.: Berkeley Linguistics Society.
Plunkett, K. & Marcham, V. (1990). From rote learning to system building. (Technical Report No. 9020). San Diego, Calif.: Center for Research in Language, University of California.
Plunkett, K. & Marchman, V. (1991). U-shaped learning and frequency effects in a multilayered perceptron: Implications for child language acquisition. *Cognition, 38*, 43–102.
Pyles, T. & Algeo, J. (1982). *The origins and development of the English language*, 3rd edn. New York: Harcourt Brace Jovanovich.
Rosenblatt, F. (1962). *Principles of neurodynamics*. New York: Spartan.
Rueckl, J.G., Cave, K.R., & Kosslyn, S.M. (1989). Why are "what" and "where" processed by separate cortical visual systems? A computational investigation. *Journal of Cognitive Neuroscience, 1*, 171–186.
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 microstructure of cognition. Vol. 2: Psychological and biological models*. Cambridge, Mass.: Bradford Books/MIT Press.
Seidenberg, M.S. & McClelland, J.L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review, 96*, 523–568.
Senghas, A., Kim, J.J., Pinker, S., & Collins, C. (1991). Plurals-inside-compounds: Morphological constraints and their implications for acquisition. Paper presented at the *Sixteenth Annual Boston University Conference on Language Development*, 18–20 October.
Slobin, D.I. (1971). On the learning of morphological rules: A reply to Palermo and Eberhart. In D.I. Slobin (Ed.), *The ontogenesis of grammar: A theoretical symposium*. New York: Academic Press.
Slobin, D.I. (1977). Language change in childhood and in history. In J. Macnamara (Ed.), *Language learning and thought*. New York: Academic Press.
Sloman, S.A. & Rumelhart, D.E. (1991). Reducing interference in distributed memories through episodic gating. In A. Healy, S.M. Kosslyn, & R.M. Shiffrin (Eds), *From*

learning processes to cognitive processes: Essays in honor of William K. Estes. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Sproat, R. (1992). *Morphology and computation*. Cambridge, Mass.: MIT Press.

Stemberger, J.P. (1990). *The acquisition of morphology: Analysis of a symbolic model of language acquisition*. Unpublished manuscript, Department of Linguistics, University of Minnesota.

Ullman, M. & Pinker, S. (1990). Why do some verbs not have a single past tense form? Paper presented to the *Fifteenth Annual Boston University Conference on Language Development*, 19–21 October.

Ullman, M. & Pinker, S. (1991). Connectionism versus symbolic rules in language: A case study. Paper presented at the *Spring Symposium Series of the American Association for Artificial Intelligence*, Stanford, Calif., 26–28 March.

Wasow, T. (1981). Comments on the paper by Baker. In C.L. Baker & J.J. McCarthy (Eds), *The logical problem of language acquisition*. Cambridge, Mass.: MIT Press.

Zipf, G. (1935/1965). *The psycho-biology of language: An introduction to dynamic philology*. Cambridge, Mass.: MIT Press.

APPENDIX 1

Stem	Vowel-change Form	Stem Rating (Exps 1 & 2)	Suffixed Rating (Exps 1 & 2)	Vowel-change Rating (Exps 1 & 2)	Suffixed Production Probability (Exp. 3)	Non-suffixed Production Probability (Exp. 3)	Stem Phonological Goodness	Suffixed Phonological Goodness	Vowel-change Phonological Goodness
Prototypical pseudo-irregular									
spling	splung	5.7	4.5	5.3	0.71	0.33	4.8	2.5	4.1
skring	skrung	4.1	4.0	4.4	0.75	0.38	5.8	4.1	5.3
sprink	sprunk	5.9	5.6	5.3	0.83	0.21	6.1	3.8	5.5
cleed	cled	5.6	6.0	5.5	0.83	0.42	5.6	3.8	5.8
preed	pred	6.0	5.3	5.6	0.63	0.25	5.5	2.3	4.6
queed	qued	5.4	4.0	4.8	0.75	0.38	4.9	3.3	5.3
cloe	cloo	6.3	6.1	3.4	0.92	0.21	5.6	4.6	6.5
froe	froo	6.0	6.0	5.4	0.92	0.63	5.5	5.1	4.9
plare	plore	5.8	6.5	4.3	0.83	0.29	5.8	5.3	5.5
quare	quore	5.9	6.6	5.8	0.96	0.46	6.1	5.4	5.4
Intermediate pseudo-irregular									
fring	frung	5.6	3.3	5.1	0.79	0.33	5.6	3.3	5.1
ning	nung	5.2	4.0	5.8	0.92	0.71	4.9	4.0	4.0
frink	frunk	5.9	5.3	4.5	0.83	0.63	5.1	3.3	5.8
cleef	clef	5.7	5.5	3.4	0.71	0.33	5.9	5.4	6.4
preek	prek	5.8	6.4	3.3	0.96	0.79	5.3	3.8	6.0
queef	quef	5.8	5.7	3.4	0.88	0.71	4.8	5.4	5.0
foa	foo	5.5	5.7	2.6	1.00	0.42	5.1	5.9	5.0
voa	voo	5.6	5.6	2.3	0.92	0.17	5.0	2.8	5.4
jare	jore	5.9	5.8	4.1	0.96	0.71	5.1	2.5	5.6
grare	grore	4.5	5.3	4.0	0.92	0.83	4.6	6.3	3.0
Distant pseudo-irregular									
trisp	trusp	5.6	5.9	4.2	0.96	0.17	6.1	4.3	6.0
nist	nust	5.9	6.0	4.1	0.88	0.21	5.8	3.8	5.8
blip	blup	6.6	7.0	3.0	0.96	0.13	6.6	3.9	4.5
gleef	glef	5.1	6.2	2.9	0.79	0.17	5.6	3.4	5.8
keeb	keb	5.6	5.6	3.9	0.92	0.21	5.0	4.1	5.8
meep	mep	6.4	6.1	2.9	0.92	0.54	6.1	4.1	5.6
goav	goov	5.0	5.5	3.0	0.96	0.25	4.4	3.5	4.9
joam	joom	5.2	5.7	2.2	0.88	0.21	3.9	2.4	6.3
flape	flope	5.4	5.9	3.7	1.00	0.21	5.5	3.8	5.5
blafe	blofe	4.9	6.1	4.2	0.96	0.38	5.3	3.5	5.1
Prototypical pseudo-regular									
plip	plup	6.2	6.1	2.9	0.92	0.38	6.0	5.1	5.3
glip	glup	6.3	6.7	2.6	1.00	0.42	6.3	2.8	5.9
brip	brup	5.5	6.0	3.7	0.96	0.67	6.0	4.5	5.0
gloke	glook	5.4	6.4	3.1	0.96	0.29	5.5	4.6	5.0
proke	prook	6.0	6.3	3.9	1.00	0.38	5.9	2.9	4.3
greem	grem	5.7	6.3	3.6	0.96	0.38	6.1	4.3	5.5
pleem	plem	5.5	5.9	3.0	1.00	0.33	4.9	3.5	5.9
treem	trem	5.7	6.0	3.7	0.83	0.17	5.8	2.8	6.1
slace	sloce	5.7	6.4	3.4	1.00	0.17	5.3	2.4	4.9
nace	noce	5.7	6.1	3.1	0.92	0.54	6.3	3.6	4.5

Appendix 1 (cont.)

Stem	*Vowel-change Form*	*Stem Rating (Exps 1 & 2)*	*Suffixed Rating (Exps 1 & 2)*	*Vowel-change Rating (Exps 1 & 2)*	*Suffixed Production Probability (Exp. 3)*	*Non-suffixed Production Probability (Exp. 3)*	*Stem Phonological Goodness*	*Suffixed Phonological Goodness*	*Vowel-change Phonological Goodness*
Intermediate pseudo-regular									
brilth	brulth	3.2	4.6	3.3	0.83	0.25	4.3	3.3	2.9
glinth	glunth	4.5	4.9	3.9	1.00	0.33	5.1	2.3	5.0
plimph	plumph	4.4	6.0	3.7	0.88	0.29	4.6	2.1	4.3
ploab	ploob	4.8	6.0	2.7	1.00	0.58	4.9	2.8	4.9
ploag	ploog	4.7	6.0	2.1	1.00	0.33	4.4	3.4	5.3
smeeb	smeb	4.7	4.1	5.6	0.88	0.54	5.1	2.1	4.9
smeeg	smeg	4.9	5.2	3.8	0.79	0.38	4.6	1.9	6.1
smeej	smej	3.6	5.2	2.8	0.96	0.50	4.5	1.6	5.0
smaib	smobe	3.9	5.0	4.6	0.92	0.13	4.4	1.9	5.0
smaig	smoag	4.3	5.2	3.2	0.96	0.08	5.8	2.4	4.8
Distant pseudo-regular									
frilg	frulg	3.9	4.7	3.5	0.96	0.79	3.1	2.9	3.4
krilg	krulg	3.5	5.5	3.0	0.88	0.38	3.8	3.3	3.1
trilb	trulb	3.7	5.4	2.8	0.92	0.17	3.8	3.4	3.4
ploamph	ploomph	2.7	5.3	3.0	1.00	0.54	2.9	3.4	3.1
ploanth	ploonth	2.8	4.4	2.9	0.88	0.46	3.1	2.5	5.0
smeelth	smelth	3.0	3.8	5.0	0.75	0.29	4.0	3.4	5.0
smeenth	smenth	3.5	4.2	4.8	0.83	0.29	3.9	2.3	5.0
smeerg	smerg	3.5	4.9	3.2	0.67	0.17	3.5	5.5	5.5
smairg	smoarg	3.2	5.7	3.9	0.96	0.75	3.3	2.3	2.9
smairph	smoarph	3.3	6.0	3.1	0.92	0.54	3.3	2.5	2.3

APPENDIX 2: WHY DO LANGUAGES HAVE IRREGULAR FORMS?

The distinct computational requirements posed by the existence of irregular pairs in a language raises the question of why languages have irregulars to begin with. In this appendix, we discuss some relevant data from our experiments and ancillary ratings.

Irregular forms have often been seen as the quintessence of linguistic illogic and quirkiness (e.g. Lederer, 1989), but Bybee (1985) notes that they may have some advantages. No language uses rule-like productivity nearly as much as it could. For example, one can imagine a language with separate morphemes indicating humanness, animacy, colour, size, number of legs, ability to fly, and so on, allowing most animals to be named by predictable productive combinations. But in all languages, such concepts are instead given independent, unanalysable labels that have to be memorised by rote. Presumably, this reflects a trade-off between productivity and processability (Cherry, 1966; Mandelbrot, 1954; Slobin, 1977; Zipf, 1935/1965). Unique strings of recurring morphemes for every distinguishable concept would be easily analysable but would necessarily be long and complex; unanalysed labels would be limited to a fixed number that have been memorised beforehand by both speaker and hearer, but can be short, simple, and recognised by simple look-up. Given that humans have the ability both to generate linguistic objects by rule and to retrieve them from memory – grammar and lexicon, respectively – nothing should prevent them from memorising simple unanalysed forms for meaning combinations that could in principle be generated by rule.

Bybee speculates that combinations of stem meaning and affix meaning that are cognitively cohesive and high in frequency are liable to have their forms listed in memory and reduced in length and complexity. The properties of existing English irregular and regular verbs are consistent with her conjecture. Pinker and Prince (1988) note that both the stem forms and the past-tense forms of irregulars are uniformly short (monosyllabic or with monosyllabic roots), conforming to the prosodically minimal words of English, which may be the canonical word shapes of the language (McCarthy and Prince, in press.) In contrast, the English suffixation operation necessarily lengthens the word and may create an unnatural phonological sequence where the word-final consonants and the suffix are concatenated.[12] Indeed, a *d* or *t* at the end of a consonant cluster appears to be fragile in perception, production or both. We have seen examples of its omission in written versions of forms that are easy to misanalyse as not being morphologically complex, such as concatenatives (*use to*, *suppose to*), clichés (*cut and dry*, *if worse comes to worse*, *there's no love loss between those two*), and adjectives (*Broil Cod*, *Use Books*, *Blacken Redfish*, *Handicap Facilities Available*, *This Station Close*). Some have actually replaced the originals in the language, such as *ice tea* and *skim milk*. (In addition, occasionally, they are subject to hypercorrection as in *colored TV* and *soured cream*.)

Table 1 and Appendix 1, which report ratings of pure phonological goodness, contain relevant data. The phonological manipulation used to create forms similar to English irregular pasts, a vowel change, had a very different effect on phonological naturalness than the manipulation used to create forms similar to English regular pasts, the addition of a terminal consonant. Changing the vowel had no effect for words in four of the six classes; in one of the classes, the forms with the vowel change were actually rated as better sounding [$F(1,23) = 5.94$, $P < 0.05$]; in one other case, they were rated worse [$F(1,23) = 16.18$, $P < 0.001$]. Adding a consonant, however, made all six significantly worse than the stem or vowel change forms ($P < 0.05$).

---

[12] Some 25 verbs with *-en* participles, and the nouns *oxen* and *children*, do have polysyllabic irregular forms, but even these are phonologically quite natural.

The fact that internal vowel changes do not alter the basicness of a word, whereas suffixation does, also helps makes sense of the fact that English irregular verbs, though consistently basic in modern English, are the result of two independent historical processes. One is the retention of vowel-change patterns from Indo-European ablaut rules ("strong" verbs). The second process implicates the target of phonological basicness more directly. Many current irregular verbs (the "weak" irregulars) appear to have resulted by the erosion of weak verbs that had been suffixed with an affix containing *t* or *d*. These changes yielded monosyllables, often with single final consonants rather than consonant clusters (e.g. *have–had*, *make–made*, *bend–bent*, *hit–hit*, *teach–taught*). Indeed, no irregular verb has the syllabic ending *-Id* seen in regular verbs that end in *t* or *d* (Pinker & Prince, 1988).

In sum, the differences in English-speakers' intuitions of phonological naturalness between vowel changes and consonant addition supports the conjecture that the co-existence of regular and irregular alternations in a language represents a trade-off between full productivity at the cost of increased phonological complexity (regulars) and input-dependent item-wise acquisition with the benefit of lexical basicness (irregulars).