Journal of Memory and Language 61 (2009) 55–76

Contents lists available at ScienceDirect

Journal of Memory and Language

journal homepage: www.elsevier.com/locate/jml

Past-tense generation from form versus meaning: Behavioural data and simulation evidence

Anna M. Woollams [a,*], Marc Joanisse [b], Karalyn Patterson [c]

[a] Neuroscience and Aphasia Research Unit, School of Psychological Sciences, Zochonis Building, University of Manchester, Brunswick Street, Manchester M13 9PL, England, UK
[b] Department of Psychology, The University of Western Ontario, London, Ontario, Canada
[c] MRC Cognition and Brain Sciences Unit, Cambridge, UK

ARTICLE INFO

Article history:
Received 28 December 2006
revision received 6 February 2009
Available online 19 March 2009

Keywords:
Past tense
Inflection
Morphology
Semantics
Frequency
Regularity

ABSTRACT

The standard task used to study inflectional processing of verbs involves presentation of the stem form from which the participant is asked to generate the past tense. This task reveals a processing disadvantage for irregular relative to regular English verbs, more pronounced for lower-frequency items. Dual- and single-mechanism theories of inflectional morphology are both able to account for this pattern; but the models diverge in their predictions concerning the magnitude of the regularity effect expected when the task involves past-tense generation from meaning. In this study, we asked normal speakers to generate the past tense from either form (verb stem) or meaning (action picture). The robust regularity effect observed in the standard form condition was no longer reliable when participants were required to generate the past tense from meaning. This outcome would appear problematic for dual-mechanism theories to the extent that they assume the process of inflection requires stem retrieval. By contrast, it supports single-mechanism models that consider stem retrieval to be task-dependent. We present a single-mechanism model of verb inflection incorporating distributed phonological and semantic representations that reproduces this task-dependent pattern.

Introduction

The mechanisms by which speakers produce the past tense of both regular and irregular verbs has proven an arena for debate between dual-mechanism theories (e.g., Pinker & Ullman, 2002; Tyler et al., 2002a) and connectionist single-mechanism approaches (e.g., Joanisse & Seidenberg, 1999; McClelland & Patterson, 2002) to some aspects of language processing. This is in part because the domain of past-tense inflection represents a paradigmatic example of the quasi-regularity that is also characteristic of a number of other domains of language, perhaps most notably – in English – the translation from orthography to phonology required in reading words aloud (Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001; Plaut, McClelland, Seidenberg, & Patterson, 1996). Quasi-regularity in the English past-tense domain refers to the combination of facts (a) that the great majority of past-tense forms are predictable, entailing only addition of the –*ed* suffix to the stem, but (b) that some verbs violate this rule, requiring either no change to the stem (e.g., *cut, hit*), or some transformation of the stem itself (e.g., *ran*, *stole*, *blew*, etc.).

To date, research into the processes underlying past-tense generation has focussed almost exclusively upon data from what we will call the Stem Inflection task. In this task, participants are given a verb stem, typically written though sometimes spoken, and required to produce that verb in its past-tense form. The latency and accuracy of the response are presumed to reflect the underlying mechanisms involved in past-tense inflection. The most robust

* Corresponding author. Fax: +44 161 275 2873.
E-mail address: anna.woollams@manchester.ac.uk (A.M. Woollams).

0749-596X/$ - see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.jml.2009.02.002

empirical finding to emerge in the Stem Inflection paradigm is a regularity effect: longer response times (RTs) and/or lower accuracy for irregular than for regular verbs, particularly for lower-frequency items (Prasada, Pinker, & Snyder, 1990; Seidenberg & Bruck, 1990). Errors to irregular verbs usually take the form of regularisations (e.g., *cutted, freezed, stealed)*, and such responses are even more prevalent amongst children who have not yet fully mastered inflectional morphology (Marchman, 1997) and in certain neurological patients as a consequence of brain damage (Cortese, Balota, Sergent-Marshall, Buckner, & Gold, 2006; Patterson, Lambon Ralph, Hodges, & McClelland, 2001; Patterson et al., 2006). These behavioural effects have invited two different theoretical accounts in the form of dual- versus single-mechanism models of inflectional morphology.

The widespread focus on performance in the Stem Inflection task has been justified by an often implicit assumption that this task reflects the same processes underlying production of the past tense in natural language (Pinker, 1999, p.109). Data from the speech of children during language acquisition (Marcus et al., 1992) and of some neurological patients (Patterson et al., 2001), however, offer reasons to question this assumption. In these populations, regularisation errors in spontaneous speech are more likely to occur when there has been very recent exposure to the verb's stem form, suggesting that the Stem Inflection task may artificially inflate the magnitude of the regularity effect, possibly as a result of phonological priming from the stem to an "add –ed" response (Marcus et al., 1992; Okrent, 2004; Ramscar, 2002; Stemberger, 2004). The aim of the present study was to determine whether the robust regularity effect observed in Stem Inflection holds when the task is to generate the past tense of a verb from an action picture, which we will refer to as Picture Inflection. Not only does this comparison across tasks speak to the debate between existing dual- and single-mechanism accounts; results of the Picture Inflection task may also provide constraints upon the future development of these models by giving evidence from a procedure more similar to inflection in normal language production.

Past-tense generation from form

Previous results from the Stem Inflection task have been regarded as basic facts about normal inflectional processing which any adequate theory must be able to explain. As noted above, the regularity effect in this task is qualified by an interaction with frequency, such that regularity effects are more pronounced for lower-frequency items, and frequency effects more pronounced for irregular items (Prasada et al., 1990; Seidenberg & Bruck, 1990). This result obtains irrespective of whether the verb stimuli vary in the frequency of only the past-tense form or the summed frequency of all of a verb's forms, otherwise known as its lemma frequency (Seidenberg, 1992).

Another influence upon performance in the Stem Inflection task is the consistency of a verb's phonological rime neighbourhood in terms of the associated present-to-past-tense transformation (Marchman, 1997). Specifically, Seidenberg and Bruck (1990) reported that inconsistent regular verbs with stems that happen to rhyme with the stems of irregular verbs (e.g., *ping-pinged* cf. *sing-sang* and *ring-rang*) are inflected more slowly and less accurately than consistent regular verbs with stems that rhyme only with the stems of other regular verbs (e.g., *rush-rushed* cf. *hush-hushed*, *gush-gushed*, etc.). Further, Bybee and Moder (1983) found similar effects of neighbourhood consistency upon nonce verb inflection in skilled adults, such that stems with inconsistent neighbourhoods (e.g., *spling* cf. *ping-pinged, sing-sang* and *ring-rang)* were sometimes inflected in accordance with their irregular counterparts (i.e. *splang*) rather than the regular transformation (i.e. *splinged*).

Dual-mechanism accounts

The central tenet of dual-mechanism theories of inflectional morphology is that successful inflection of irregular versus regular/nonce verbs requires two functionally independent procedures, underpinned by distinct neural substrates (Pinker & Ullman, 2002; Tyler, Randall, & Marslen-Wilson, 2002b; Tyler et al., 2002a; Ullman, 2001; Ullman et al., 1997; Ullman et al., 2005). The strongest form of this view is provided by the Words and Rules theory (Pinker 1991; Pinker 1994; Pinker & Prince, 1988), and we shall therefore focus on this particular account. This perspective derives from traditional linguistic theory, in which the lexicon and grammar are independent cognitive subsystems. In this framework, all monomorphemic words, including verb stems, have entries in the lexicon. Input to the inflectional system consists of the word stem (e.g. *walk* or *hold*) plus the grammatical feature of past tense (Pinker & Ullman, 2002, p.457). If a verb is irregular (*hold*), then its past-tense form (*held*) will also be listed in the lexicon. In at least the original formulation of this account (Pinker 1991; Pinker 1994; Pinker & Prince, 1988), if a verb is regular (*walk*), then its past-tense form (*walked*) will not be listed, but rather will be created when required through application of the generalised grammar that encodes the regular past-tense rule of "add –ed". An irregular verb escapes regularisation (*holded*) by virtue of the presence of its correct past-tense form in the lexicon, retrieval of which blocks application of the rule.

Within the Words and Rules account, the fact that the rule-blocking process necessary for correct inflection of irregular verbs takes time, and may occasionally fail, provides an explanation for the regularity effect obtained in the Stem Inflection task. The principle that efficiency of retrieval from the lexicon is modulated by word frequency accounts for its observed interaction with regularity. Specifically, the representations of the stems of high-frequency irregular verbs will be activated rapidly, as will their correct past-tense forms. This rapid retrieval allows for efficient blocking of the regular rule, and hence there is little cost for high-frequency irregulars relative to regulars. Similarly, as only irregular past-tense forms must be retrieved from the lexicon, a significant impact of the frequency of past-tense forms will be confined to these items, although the frequency of the stem is presumed to affect processing of irregular and regular verbs alike (Pinker, 1999).

The effects of the present-to-past-tense consistency of a verb stem's phonological neighbourhood were not predicted by the original form of the Words and Rules theory. In later incarnations (Pinker, 1999; Pinker & Ullman, 2002; Ullman et al. 1997; Ullman et al. 2005) however, it has been suggested that there is partial sublexical activation of verb stems' phonological neighbours within the lexicon, entailing a danger of "irregularisation" of inconsistent regular verbs (e.g., *ping-pang*). To overcome this problem, it has been proposed that the regular past-tense forms of such items are also stored in the lexicon, and hence these items are processed more slowly as they are subject to the same blocking mechanism that applies to the irregular forms. Of course, the regular past-tense forms of nonce verbs cannot be stored in the lexicon, by definition, hence the sublexical activation of verb stems' phonological neighbours may result in an analogical inflection (e.g., *spling-splang).* With these modifications, the dual-mechanism approach is in principle able to account for the previously outlined basic facts about normal past-tense generation derived from the standard Stem Inflection task.

Single-mechanism accounts

Single-mechanism models of inflectional morphology propose that one set of procedures is sufficient to allow correct processing of irregular and regular/nonce forms (Daugherty & Seidenberg, 1992; Daugherty & Seidenberg, 1994; Joanisse & Seidenberg, 1999; McClelland & Patterson, 2002; Rumelhart & McClelland, 1986). This framework derives from a broader connectionist approach to human information processing, in which transformations between forms are accomplished according to multiple probabilistic mappings learnt from exposure. In terms of the past tense, this framework proposes that inflection proceeds with reference to general semantic and phonological knowledge for all verbs, irrespective of their regularity or familiarity. Processing within computational models that implement this single mechanism is determined by the weights on connections between representational units that reflect the distribution of mappings present in the set of items on which they are trained. Stem Inflection for regular verbs receives substantial community support, as the mappings that preserve the stem form and add an allomorph of the –ed suffix are overwhelmingly the most prevalent within the training corpus. In contrast, Stem Inflection for irregular verbs receives restricted community support due to the relative rarity of the mappings involved, and also suffers interference from the more common regular mappings.

The fact that successful inflection of irregular items requires inhibition of competition, which takes time and may occasionally fail, provides single-mechanism models with an explanation of the regularity effect observed in the standard Stem Inflection task. Another major determinant of the strength of connections within such accounts is the frequency of individual items, and this aspect of single-mechanism models accounts for the interaction between frequency and regularity. Only common irregular verbs, with which the network has much experience, can overcome interference from the regular mapping efficiently, counteracting a regularity effect for these items. Low-frequency regular verbs benefit from the weights that encode the predominance of this mapping across the network's vocabulary, counteracting a frequency effect for these items.

The effects of the past-tense consistency of a verb stem's phonological neighbourhood seen in the Stem Inflection task represent a direct prediction of the single-mechanism approach (Seidenberg, 1992). Given that processing within single-mechanism models is largely determined by the degree of consistency of a particular verb within the context of the phonological rime neighbourhood of its stem, it falls naturally from this account that regular inconsistent verbs (e.g., *ping-pinged* cf. *sing-sang* and *ring-rang*) will suffer interference from the mappings associated with their irregular verb neighbours, which will slow correct inflection relative to regular consistent verbs (e.g., *rushed-rushed* cf. *hush-hushed*, *gush-gushed*, etc.), and may occasionally result in an irregularisation error (e.g., *ping-pang).* Early single-pathway implementations of the single-mechanism approach, designed to model the Stem Inflection task (Daugherty & Seidenberg, 1992; Daugherty & Seidenberg, 1994; Seidenberg, 1993), have successfully simulated both the regularity-by-frequency interaction and the consistency effect observed for regular verbs.

Exactly the same logic applies to the inflection of nonce forms, such that a verb stem from an inconsistent neighbourhood (e.g., *spling* cf. *ping-pinged, sing-sang* and *ring-rang*) will occasionally be inflected according to the mappings associated with its irregular neighbours, rather than the regular mappings that predominate across the whole corpus. In fact, Ramscar (2002) has demonstrated that the likelihood of an irregular inflection for a nonce verb may be altered according to the semantic context in which is presented (e.g., *splink-splank* is more likely in contexts favouring a meaning associated with drinking than blinking). Such results accord with the single-mechanism assumption that inflection proceeds with reference to both general phonological and semantic knowledge. Moreover, they demonstrate the flexibility of the speaker's relative reliance on these two sources of knowledge in service of the task at hand.

Past-tense generation from meaning

As already noted, many researchers assume that the Stem Inflection task is a good analogue of past-tense production in natural language. Pinker (1999) makes this assumption explicit when he refers to the Stem Inflection task as a situation where people must "cough up past-tense forms under time pressure" and follows this immediately with "as they do in rapid conversation" (p. 129). The primary motivation for the current study was to evaluate the validity of this assumption. Effectively, this question reduces to whether production of the past tense is obligatorily preceded by retrieval of the verb stem. If stem retrieval is necessary, then patterns of performance in past-tense production should be largely the same whether the participant starts with the verb's stem or with meaning. Although considerations of past-tense generation from meaning have been rare in the inflectional morphology literature, one relevant source of such evidence is corpus studies of inflectional errors in natural speech production.

A comparison of the regularisation error rates in connected speech to those obtained in the Stem Inflection task certainly suggests some important differences. The rate of such errors observed in the conversational speech of normal adult speakers has been estimated at .00004% (Marcus et al., 1992, p. 45), which is vanishingly small compared to the 2.6% regularisation rate reported by Seidenberg and Bruck (1990) in Stem Inflection. Similarly, the regularisation error rate observed in a meta-analysis of children's spontaneous speech was 2.5% (Marcus et al., 1992, p. 35), as compared to over 10% (Marchman, 1997) in a Stem Inflection task involving a sentential context (e.g., "Every day I dig a hole. Yesterday I ___ a hole."). In addition, patients with semantic dementia, who are very prone to regularisation errors in the Stem Inflection task (Cortese et al., 2006; Patterson et al., 2001; Patterson et al., 2006), do not appear to make regularisation errors in spontaneous speech with any greater frequency than normal speakers do.

Although these discrepancies are striking, it is of course possible that they arise for reasons other than presentation of the verb stem in one case and not in the other. Critically, different sets of items are likely to enter into the two analyses. Lower-frequency verbs are not often used in conversation but are favoured by the designers of experimental tasks. Given that the regularity effect in Stem Inflection is modulated by frequency, the likely consequence of the bias towards lower-familiarity verbs in this task is to inflate the regularisation rate relative to spontaneous speech. Children and patients with semantic dementia, who may well not know the full meaning of lower-frequency verbs, are even less likely than normal adults to include these words in their spontaneous speech; but this does not prevent experimenters from confronting individuals from these populations with such items in Stem Inflection. Hence these comparisons can at best provide only hints toward answering the question of interest here.

Marcus et al. (1992, p. 64) reported an analysis that suggests more compelling evidence against the idea of mandatory stem retrieval prior to past-tense production in connected speech. These authors observed that Abe, a child originally studied by Kuczaj (1976), was significantly more likely to regularise a particular verb if his parents had used the stem form in the preceding conversational turn (e.g., Mother: "And what did you choose to do?"; Abe: "I choosed to make cookies."), despite correct inflection of the same verb on other occasions. This pattern was also observed in the speech of two of the three children reported by Brown (1973). As these regularisation error rates were calculated across the same items, the difference cannot be attributed to a lack of knowledge of the verb's meaning or its correct past-tense form. Instead, the result was attributed to priming of the child's representation of the stem form through its use by the parent, increasing the likelihood of regularisation. A similar phenomenon has been reported by Patterson et al. (2001) for semantic dementia: the rare instances of regularisations in the conversational speech of such patients seem to be limited to occasions when the patient has just heard the stem form (e.g., Experimenter: "Did you fly directly to Australia?"; DM: "Yes, I flied there.").

An experimental assessment of the extent to which stem retrieval is mandatory in past-tense production will require the same set of items to be employed in a contrast between the Stem Inflection task and a generation from meaning task that does not involve presentation of the stem. We are aware of only one previous attempt to achieve this contrast. Okrent (2004) compared performance by adult speakers in the standard Stem Inflection task to that in a past-tense generation task involving sentence contexts designed to elicit the target verb without presentation of its stem. In the generation from meaning task, an initial context sentence (e.g., "There was one more place at the conference table.") was presented at the top of the screen for 2000 ms; followed by a second context clause on the line below (e.g., "He went to the chair, and") for a duration corresponding to 120 ms per syllable; followed by the final clause on the line below (e.g., "he ____ down."), which remained on screen until the participant responded. Okrent (2004) reported a significant regularity effect in both RTs and errors in the Stem Inflection task but not in the sentence-completion condition. The regularity-by-frequency interaction in the Stem Inflection task was, however, the reverse of that usually observed, with the regularity effect larger for high- than for low-frequency verbs. In addition, timing of the response from the onset of the final clause of the sentence in the sentence-completion condition may have increased overall RTs and reduced sensitivity to a difference between regular and irregular verbs. Thus, although the results of this pioneering study seem to argue against the notion of mandatory stem retrieval in past-tense generation from meaning, additional evidence is required.

We set out to provide such evidence in the present study, where we examined the patterns of performance in the standard Stem Inflection task and in a novel Picture Inflection task that required participants to generate the past tense from depictions of the same verbs in the form of line drawings of actions. Consideration of the results obtained when past-tense generation proceeds from form versus meaning provides an explicit assessment of the widely held assumption that the former is a good analogue of the latter. In addition, we believe that dual- and single-mechanism models of inflectional morphology make different predictions concerning performance in the two tasks.

Dual-mechanism predictions

Accounts of inflectional morphology are, of course, ultimately intended to explain not only performance in laboratory tasks but also the mechanisms underlying past-tense processing in normal language comprehension and production. As noted earlier, fundamental to dual-mechanism theories like Words and Rules is the distinction between lexicon and grammar, and descriptions of inflectional processing typically begin with retrieval of the stem from the lexicon before application or blocking of the grammatical rule. These accounts rarely if ever seem to address the role of semantic information that must logically underlie past-tense production from meaning, and as a consequence "They have been vague about whether the basic phonological processing of the base form takes place first as a necessary step on the way towards generating the irregular form." (Stemberger, 2004, p. 83). Yet this issue is clearly critical when it comes to deriving predictions concerning the performance expected when participants are required to generate the past tense from meaning. Given Pinker's (1999, p. 129) previously cited assumption of comparability between generation of past tense in Stem Inflection and conversation, we can only surmise that activation of the verb stem represents a common feature of both situations.

If this surmise is correct, then the inflectional processes responsible for producing the past tense should be largely impervious to the nature of the input. Whenever an inflected form is required, the morphological system will receive a demand for the desired grammatical form, e.g., plural (if we are talking about two shops) or progressive (if we are currently shopping) or past tense (if we shopped yesterday) etc., from the semantic system. With respect to past-tense generation from form versus meaning, the input constitutes the externally presented verb stem in the former case and the internally generated semantic representation of the verb in the latter.

The dual-mechanism approach apparently assumes that production of the past tense proceeds in the same fashion in the two tasks, as it is only after the lexical representation of the stem has been activated that the past tense can either be recovered from the lexicon, for an irregular verb, or created by application of the grammatical rule, for a regular verb. Dual-mechanism models therefore seem to predict that the regularity effect that characterises performance in Stem Inflection should also emerge when the task is past tense generation from meaning, as in Picture Inflection.

Single-mechanism predictions

There have been several previous implementations of single-mechanism computational models specifically intended to simulate data from the standard Stem Inflection task (Daugherty & Seidenberg, 1992; Daugherty & Seidenberg, 1994; Rumelhart & McClelland, 1986). Although inflection within such models involves a verb stem's phonology as input, this is by no means a theoretical commitment of the approach, and later models have turned their attention to the role of semantic information in activating the phonological representations of the simple past tense. Hoeffner (1992; Hoeffner & McClelland, 1993) presented a single-mechanism model of the mapping of semantics to phonology for regular and irregular verbs (see Ramscar and Yarlett (2007) for a similar approach to plurals). This model was trained to generate all aspects of an inflectional paradigm (e.g., *jump, jumps, jumping, jumped*) from an abstract semantic representation, without any need for intermediate stem activation. Simulation of performance in Stem Inflection, which involved activation of the stem's phonology as input, revealed the standard regularity-by-frequency interaction; but performance in past-tense generation directly from semantics was not assessed, and hence the relative magnitude of the regularity effect obtained without external activation of stem phonology remains unknown.

Further consideration of the role of semantic information in inflection is provided by a large scale implementation presented by Joanisse and Seidenberg (1999). In this model, hidden units linking distributed input and output phonological representations were fully bi-directionally connected to a system of localist semantic representations. This allowed for training on a number of language production tasks over and above simple past-tense transformation from stem, including repetition, comprehension and production of both stem and past-tense forms. This model therefore illustrates the overall framework within which we can consider single-mechanism predictions concerning inflection from form versus meaning. Within this model, inflection of all verbs proceeds with reference to a combination of phonological and semantic knowledge, and what varies according to the nature of the task will be the relative balance between the two sources of information.

From the single-mechanism perspective instantiated in the Joanisse and Seidenberg (1999) model, Stem Inflection necessarily involves activation of the (input) phonological representation of the verb stem combined with activation of the semantic representation for past tense to allow production of the correct (output) phonological form. Performance is primarily determined by the frequency and predictability of the mappings between input and output phonology, which vary from typical, as is usually the case for regular verbs, to atypical, as is usually the case for irregular verbs. In this model, the presentation of the verb stem can also activate the semantic representation for that verb, which in turn can provide additional activation of the phonological past-tense form. The degree of support provided by this additional semantic activation during Stem Inflection will vary according to difficulty of the input-output mapping. For regular verbs, this mapping is highly predictable and therefore efficient, leaving little opportunity for meaning level activation to have much influence. For irregular verbs, however, this mapping is more unpredictable and difficult, and thus their processing is bolstered by the additional activation of phonology via semantics.

Of course, relative to past-tense generation from form, past-tense generation from meaning necessarily entails a greater reliance on semantic information and a lesser reliance on (input) phonological information. According to the single-mechanism view, production of the past tense from meaning entails concurrent activation of the semantic representation for a given action, such as *walking* or *running*, and the concept of that action having occurred in the past, which then activates the phonological form of the appropriate past tense, namely *walked* or *ran*. As there is no necessity for intermediate retrieval of the phonological form of the stem, the production of inflected forms from meaning within the single-mechanism framework is a "one-step" procedure (Stemberger, 2004), consistent with connectionist models of speech production (e.g., Dell, Juliano, & Govindjee, 1993; Plaut & Kello, 1999). Hence, single-mechanism models appear to predict that the disadvantage seen for irregular verbs in Stem Inflection should be reduced or even eliminated in Picture Inflection.

Experiment 1

The aim of the present research was to assess the comparability of the inflectional processes involved when participants are required to generate the past tense from form versus meaning. To address this issue experimentally, we used the same set of regular and irregular verbs in the standard Stem Inflection task and a novel task requiring inflection from a line drawing depicting an action (Picture Inflection). As outlined above, the dual-mechanism approach would seem to predict a regularity effect of comparable size in both tasks, whereas the single-mechanism approach suggests that the regularity effect obtained in Stem Inflection should be significantly reduced when the task is changed to Picture Inflection

Method

Participants

Twenty-four members of the MRC Cognition and Brain Sciences Unit volunteer panel were paid for their participation in this study; 12 of them completed the Stem Inflection task and the remaining 12 completed the Picture Inflection task. All were native speakers of British English, aged between 18 and 40 years

Stimuli

The stimuli consisted of 30 regular and 30 irregular verbs. Within each class of verbs, half were higher frequency (mean lemma frequency = 293, minimum = 92 occurrences per million) and half were lower frequency (mean lemma frequency = 46, maximum = 88 occurrences per million). All verbs were monosyllabic in both their present- and past-tense forms. These 60 items (see Appendix A) were presented as either verb stems or action pictures. The action pictures were drawn from those of Szekely et al. (2004). Properties of the regular and irregular verb stem and action picture sets are presented in Table 1. All lexical statistics are derived from the CELEX database (Baayen, Piepenbrock, & van Rijn, 1993) and values for the action pictures are based on the norms for the International Picture Naming Project (IPNP) provided by Szekely et al. (2004). An alpha level of .05 was used for all statistical tests of the stimulus properties.

The present- and past-tense frequencies were obtained from the form frequency listing of CELEX (EFW.CD), which contains multiple entries for different parts of speech. To obtain the present-tense or stem frequencies, we summed across all four instances listed for a given form (e.g., *to sleep, I sleep, you sleep, they sleep*). To calculate the past-tense frequencies, we summed across four instances listed for a given form (e.g., *I slept, you slept, he/she slept, they slept*), but did not include the frequency given for the past participle (e.g., *have slept*). For analysis, the log 10 of each verb's frequency was computed in order to normalise the distribution. A series of 2 (frequency) by 2 (regularity) between-items ANOVAs confirmed a significant main effect of frequency for: word-form frequency of the present tense, $F(1,56) = 11.71$, $p < .01$, word-form frequency of the past tense, $F(1,56) = 4.50$, $p < .01$, and lemma frequency of the verb, $F(1,56) = 19.16$, $p < .01$, with no other significant effects obtained.

To quantify the present-to-past-tense consistency of each verb, we created a ratio analogous to that presented by Ziegler, Montant, and Jacobs (1997) for reading aloud, calculated across all orthographic body-phonological rime neighbours of a verb's present tense as: *friends/(friends + enemies),* with each item included as a friend of itself (Jared, 1997). Irregular verbs had larger body neighbourhoods than regular verbs, $F(1,56) = 4.57$, $p = .04$, but no other significant effects were obtained. Consistency was calculated using both types and tokens. A significant main effect of regularity was obtained for the type present-to-past-tense consistency ratio, $F(1,56) = 571.11$, $p < .01$, with no other significant effects obtained. This indicates that, in terms of the number of neighbours, the regular verbs were of higher consistency than the irregular verbs to the same extent across high- and low-frequency items.

A significant main effect of regularity was also obtained for the token present-to-past-tense token consistency ratio, $F(1,56) = 25.37$, $p < .01$), however this was qualified by an interaction between frequency and regularity,

Table 1
Average values (and standard deviations) of a range of psycholinguistic variables as a function of regularity and frequency for the verb stems and action pictures.

Variable	Low-frequency		High-frequency	
	Irregular	Regular	Irregular	Regular
Lemma Frequency/million	52 (23)	40 (24)	295 (302)	291 (315)
Frequency/million Present	14 (7)	10 (7)	116 (131)	142 (191)
Frequency/million Past	16 (8)	14 (10)	78 (57)	76 (78)
Letter length present	4.53 (0.74)	4.27 (0.88)	4.27 (0.88)	4.27 (0.59)
Letter length past	4.40 (0.91)	6.00 (0.93)	4.40 (1.18)	6.07 (0.59)
Phonemic length present	3.73 (0.70)	3.20 (0.68)	3.33 (0.62)	3.13 (0.35)
Phonemic length past	3.80 (0.77)	4.20 (0.68)	3.40 (0.74)	4.13 (0.35)
Stem body neighbourhood	9.93 (4.38)	6.87 (6.32)	9.40 (3.70)	7.07 (4.79)
Consistency type	0.286 (0.162)	0.987 (0.052)	0.239 (0.145)	0.968 (0.060)
Consistency token	0.497 (0.345)	0.976 (0.089)	0.715 (0.242)	0.894 (0.250)
Name agreement	0.797 (0.225)	0.776 (0.277)	0.770 (0.236)	0.733 (0.233)
Name entropy	0.898 (0.779)	0.705 (0.559)	1.024 (0.779)	1.105 (0.833)
Objective complexity	21,784 (6748)	23,541 (5305)	25,110 (7224)	28,674 (7736)
Conceptual complexity	2.53 (0.74)	2.93 (0.80)	2.73 (0.70)	2.87 (0.92)
Picture naming latency	1226 (372)	1170 (192)	1223 (232)	1256 (311)

$F(1,56) = 5.24$, $p = .03$). This interaction reflects the relatively greater consistency of the high-frequency irregular items than the low-frequency irregular items in terms of the summed past-tense form frequencies of neighbours. This difference is a consequence of treating each item as a friend of itself, which inflates token consistency in proportion to the frequency of that item.

Conditions did not differ in number of letters in the stem, but in the past tense, the regular items on average had more letters than the irregular verbs, $F(1,56) = 45.46$, $p < .01$. When length is measured in phonemes rather than letters, irregular verb stems were longer than their regular counterparts, $F(1,56) = 5.54$, $p = .02$, whereas regular past-tense forms were longer than their irregular mates, $F(1,56) = 11.18$, $p < .01$.

With respect to properties of the action pictures, conditions did not differ on name agreement (% of IPNP participants producing the target name), name entropy (a measure of response agreement that takes into consideration the proportion of participants producing each alternative name) or conceptual complexity (the number of objects, animals or persons depicted in each stimulus). The regular items had higher values than the irregular items in terms of objective visual complexity, $F(1,56) = 5.78$, $p = .02$, a measure based on the size of the digitized stimuli picture files in JPEG format (see http://crl.ucsd.edu/~aszekely/ipnp/method/getdata/uspnovariables.html for details). An analysis of picture naming latencies derived from the IPNP norms, however, revealed no significant differences between conditions. The absence of a frequency effect in picture naming RTs for these stimuli is concordant with the small but significant reversed frequency effect observed for the full set of IPNP action pictures by Szekely et al. (2005). This issue and its implications for the present study will be discussed in more detail after presentation of the results of Experiment 1.

Procedure

Instructions and stimuli were displayed and RTs recorded using DMDX experimental software (Forster & Forster, 2003) using a voice-key plus headset connected to an IBM compatible Pentium II computer with a 9.99 ms refresh rate at 1024 by 768 pixel screen resolution. Erroneous responses and measurement errors resulting from inaccurate voice-key activation were recorded manually. Participants were presented with the 60 verbs as either stems or pictures, with stimulus type assigned in a counterbalanced order.

For both the Stem Inflection and Picture Inflection tasks, participants were asked to generate the simple form of the past tense aloud as soon as possible without making errors. A series of eight representative verb stems or action pictures served as practice stimuli, followed by the 60 test stimuli, presented in a different random order to each participant. Each trial began with a white cross on a black screen for 500 ms, followed by central presentation of either the verb stem or action picture. In both cases, the stimulus remained on the screen until a response was recorded or for a maximum of 4000 ms. Each vocal response was recorded from the onset of the trial for a period of 2000 ms after the voice-key trigger. A new trial then began after an interval of 420 ms.

Results

Any trial resulting in an incorrect response or inappropriate activation of the voice-key was excluded from the analysis of RTs. Error types included (1) no response prior to time-out; (2) production of the verb's past participle instead of its past tense (e.g., *steal* → "stolen" rather than "stole"); (3) re-production of the stem without 'adornment' or change; (4) analogical inflection (LARC);[1] (5) regularisations (in the case of the irregular verbs); and (6) "other name" errors. In the Stem Inflection task, this last category resulted from a visual reading error leading to production of another past-tense verb (e.g., if the word *carve* was incorrectly recognised as *crave*, then the participant would respond "craved" rather than "carved"). In the Picture Inflection task, this category of error resulted from some ambiguity concerning the intended referent leading to production of another past-tense verb (e.g., if a picture of the transaction meant to represent *selling* was interpreted as *buying*, then the participant would respond "bought" rather than "sold"). Item RT and error rate data for all conditions appear online as Supplementary material, with a breakdown of the different error types provided in Appendix B. For all subsequent analyses, an effect was considered reliable if it was significant at an alpha level of .05 by both participants (F_1) and items (F_2).

Stem Inflection

Mean correct RTs and percentage of errors of interest (regularisations plus LARC errors) for this task are presented in Figs. 1A and 2A, respectively, with exact values provided online as Supplementary material. These data were analysed with 2 (frequency) by 2 (regularity) within-participants and between-items ANOVAs. High-frequency verbs were inflected significantly more rapidly and accurately than low-frequency verbs, $F_1(1,11) = 7.49$, $p = .02$; $F_2(1,56) = 4.55$, $p = .04$; $\min F'(1,50) = 2.83$, $p = .10$; $F_1(1,11) = 7.24$, $p = .02$; $F_2(1,56) = 7.56$, $p = .01$; $\min F'(1,36) = 3.70$, $p = .06$. Regular verbs were inflected significantly more rapidly and accurately than irregular verbs, $F_1(1,11) = 17.86$, $p < .01$; $F_2(1,56) = 9.14$, $p < .01$; $\min F'(1,55) = 6.05$, $p = .02$; $F_1(1,11) = 16.04$, $p < .01$; $F_2(1,56) = 11.29$, $p < .01$; $\min F'(1,46) = 6.63$, $p = .01$. The interaction between frequency and regularity was not reliable in latency, Fs < 1, but was reliable in error rates, $F_1(1,11) = 5.04$, $p = .05$; $F_2(1,56) = 4.57$, $p = .03$; $\min F'(1,39) = 2.40$, *ns*. Consideration of simple effects revealed that the regularity effect seen in error rates was only significant for low-frequency verbs, $F_1(1,11) = 11.88$, $p = .01$; $F_2(1,56) = 15.12$, $p < .01$; $\min F'(1,31) = 6.63$, $p = .02$, and

[1] Legitimate Alternative Rendering of Components (Patterson et al., 2006), a form of error encompassing traditional regularisation responses but also including incorrect inflections by analogy to other known past tense forms (e.g., *sting* → "stang"). Such errors are not confined to irregular items, but may also occur when a regular item is inflected according to the mapping contained in an irregular neighbour (e.g., *sneeze* → "snoze").

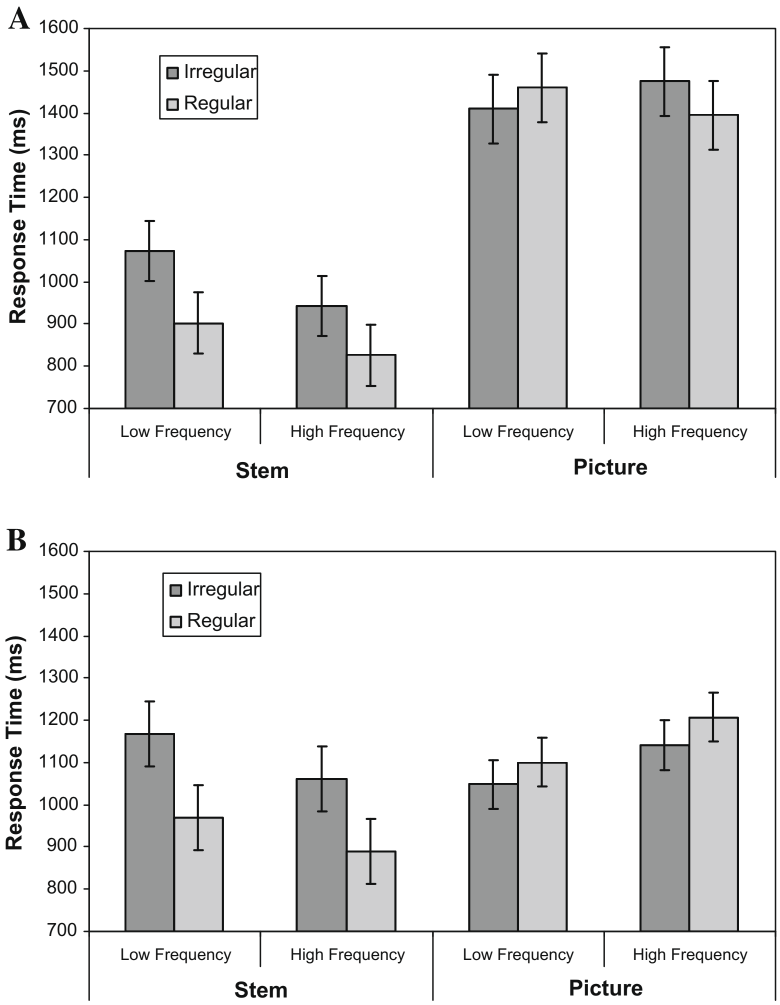

Fig. 1. Mean correct RTs as a function of frequency and regularity for Stem Inflection and Picture Inflection: (A) Experiment 1 (without familiarisation); and (B) Experiment 2 (with familiarisation). Error bars represent ±95% within-participant confidence intervals for each task (Masson & Loftus, 2003, Eq. (7), p. 212).

the frequency effect was only significant for irregular verbs, $F_1(1,11) = 6.77$, $p = .03$; $F_2(1,56) = 11.95$, $p < .01$; $\min F'(1,25) = 4.32$, $p = .05$.

Picture Inflection

Mean correct RTs and percentage of errors of interest (regularisations plus LARC errors) for this task are presented in Figs. 1A and 2A, respectively, with exact values provided online as Supplementary material. These data were analysed in the same fashion as the Stem Inflection results. In both RTs and error rates, there were no significant effects of frequency, Fs < 1; Fs < 1, or of regularity, Fs < 1; $F_1(1,11) = 1.94$, *ns*; $F_2(1,56) = 3.15$, $p = .08$; $\min F'(1,27) = 1.20$, *ns*, nor any reliable interaction between them, $F_1(1,11) = 4.97$, $p = .05$; $F_2 < 1$; $\min F' < 1$; Fs < 1.

Stem versus Picture Inflection

To provide an assessment of task differences, these data were analysed with 2 (frequency) by 2 (regularity) by 2 (stimulus type) ANOVAs. As before, regularity and frequency were treated as within-participants and between-items factors, while stimulus type was treated as a between-participants and within-items factor. With respect to RTs, Stem Inflection was significantly faster than Picture Inflection, $F_1(1,22) = 24.58$, $p < .01$; $F_2(1,56) = 134.12$, $p < .01$; $\min F'(1,30) = 20.77$, $p < .01$. Although the frequency effect was significant for Stem Inflection but not Picture Inflection, the interaction between stimulus type and frequency was not reliable, $F_1(1,22) = 2.69$, *ns*; $F_2(1,56) = 2.15$, *ns*; $\min F'(1,69) = 1.19$, *ns*. Critically, the reduction in the regularity effect seen for Picture Inflection relative to Stem Inflection was significant by participants and marginally so by items, $F_1(1,22) = 6.33$, $p = .02$; $F_2(1,56) = 2.87$, $p = .10$; $\min F'(1,78) = 1.97$, *ns*. Given that neither task yielded a significant frequency-by-regularity interaction in RTs, it is not surprising that the three-way interaction with stimulus type was not reliable, $F_1(1,22) = 4.27$, $p = .05$; $F_2 < 1$; $\min F' < 1$.

Turning to errors, there were significantly more regularisation plus LARC errors in Stem Inflection than Picture

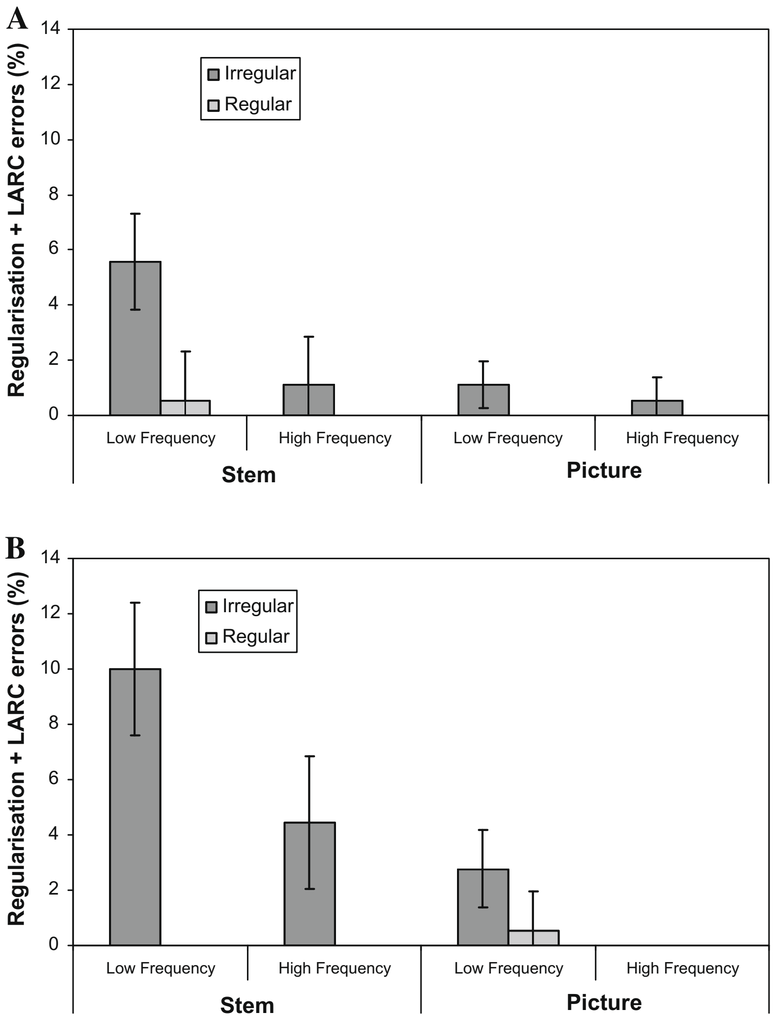

Fig. 2. Mean percentage combined regularisation plus LARC error rates as a function of frequency and regularity for Stem Inflection and Picture Inflection: (A) Experiment 1 (without familiarisation); and (B) Experiment 2 (with familiarisation). Error bars represent ±95% within-participant confidence intervals for each task (Masson & Loftus, 2003, Eq. (7), p. 212).

Inflection, $F_1(1,22) = 4.78$, $p = .04$; $F_2(1,56) = 8.05$, $p = .01$; $\min F'(1,49) = 3.00$, $p = .09$, and the frequency effect in these error types was significantly larger for Stem Inflection than Picture Inflection, $F_1(1,22) = 5.25$, $p = .03$; $F_2(1,56) = 5.15$, $p = .03$; $\min F'(1,64) = 2.60$, *ns*. The reduction in the regularity effect for Picture Inflection relative to Stem Inflection was reliable, $F_1(1,22) = 5.25$, $p = .03$; $F_2(1,56) = 5.15$, $p = .03$; $\min F'(1,64) = 2.60$, *ns*. The elimination of the regularity effect for the low-frequency items in Picture Inflection relative to Stem Inflection resulted in three-way interaction with stimulus type that approached significance by both participants and items, $F_1(1,22) = 3.37$, $p = .08$; $F_2(1,56) = 2.90$, $p = .09$; $\min F'(1,67) = 1.56$, *ns*.

Before discussing the implications of the present results, we should mention the potentially relevant task differences apparent in the "other name" error rates. As can be seen in Fig. 3A, there were significantly more other name responses for Picture Inflection than for Stem Inflection, $F_1(1,22) = 117.50$, $p < .01$; $F_2(1,56) = 36.15$, $p < .01$; $\min F'(1,77) = 27.64$, $p < .01$. Although the incidence of such responses was not affected by frequency, $F_1(1,22) = 1.71$, *ns*; $F_2 < 1$, they were reliably more common for pictures associated with regular than irregular verbs, $F_1(1,22) = 13.22$, $p < .01$; $F_2(1,56) = 4.52$, $p = .04$; $\min F'(1,78) = 3.37$, $p = .07$. The size of this inverse regularity effect did not differ for high- and low-frequency items, Fs < 1.

Discussion

With respect to the contrasting predictions derived from dual- and single-mechanism models of inflectional morphology, the results of Experiment 1 provide a clear outcome. In both RTs and regularisation/LARC error rates, the robust regularity effects observed in standard Stem Inflection were significantly reduced in Picture Inflection, with no reliable effect apparent in the latter task. This outcome appears to conflict with the predictions of dual-mechanism models. Following recognition of the picture (which of course takes more time than recogni-

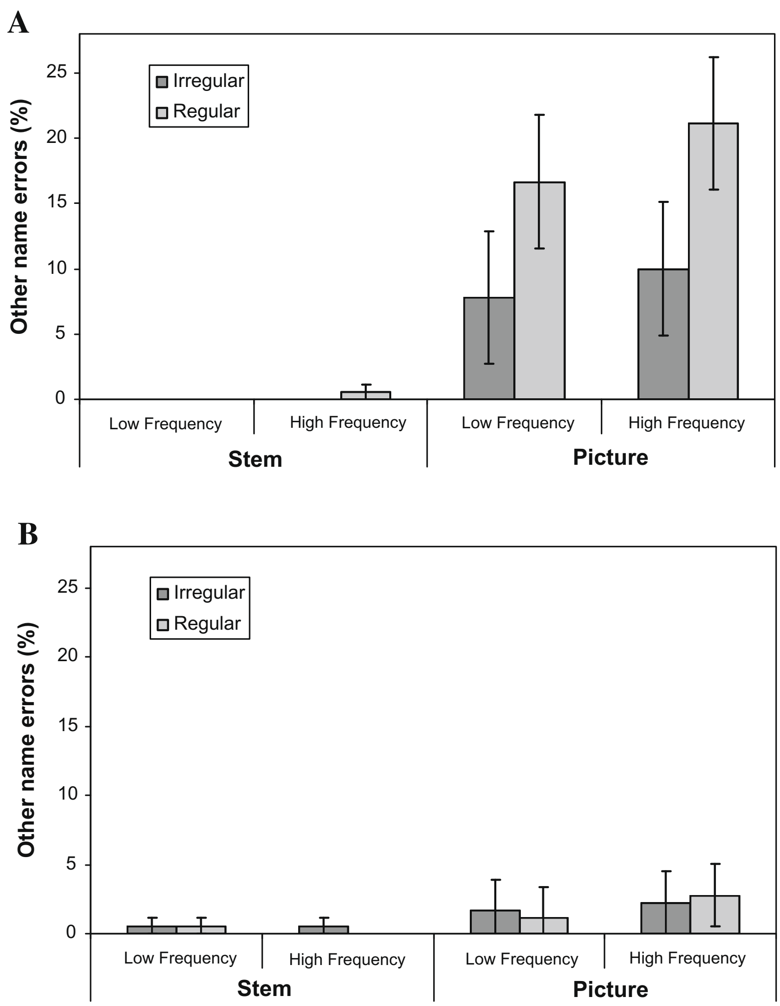

Fig. 3. Mean percentage other name error rates as a function of frequency and regularity for Stem Inflection and Picture Inflection: (A) Experiment 1 (without familiarisation); and (B) Experiment 2 (with familiarisation). Error bars represent ±95% within-participant confidence intervals for each task (Masson & Loftus, 2003, Eq. (7), p. 212).

tion of a printed word and is responsible for the overall RT difference between tasks), the dual-mechanism theory as described by Pinker (1999) assumes that the next step in the Picture Inflection task is retrieval of the verb stem from the lexicon. From that point on, the task should proceed exactly as it does in Stem Inflection, and thus should yield the standard disadvantage for irregular verbs. In contrast, the present pattern fits the predictions of single-mechanism models, which do not assume any requirement for stem retrieval in production of past tense from meaning.

Despite this clear outcome, there were some aspects of the data that warrant further comment. One somewhat unexpected result from the Stem Inflection task was the lack of a significant regularity-by-frequency interaction in RTs (cf. Prasada et al., 1990; Seidenberg & Bruck, 1990), although this interaction was apparent in regularisation/LARC error rates. We suspect that our attempt to match the high- and low-frequency irregular items for degree of inconsistency was responsible for the absence of this usual RT interaction in Stem Inflection. The present study used a present-to-past-tense consistency ratio analogous to one widely used in the study of reading aloud (Ziegler, Montant, & Jacobs, 1997), such that for each verb, the number or summed frequencies of all rhyming monosyllabic verbs characterised by the same present-to-past-tense transformation (friends) was divided by the number or summed frequencies of all rhyming monosyllabic verb stems (friends + enemies).[2] The comparable regularity effects observed for RTs to high- and low-frequency verbs may reflect the fact that the two groups of irregular verbs were matched in degree of *type* inconsistency. In contrast, the regularity-by-frequency interaction observed in regularisation/LARC error rates may have resulted from the greater

[2] For example, the irregular verb *drink-drank* has the irregular friends *shrink-shrank*, *sink-sank* and *stink-stank*; the irregular enemies *think-thought* and *slink-slunk*; and the regular enemies *blink-blinked*, *chink-chinked*, *clink-clinked*, *ink-inked*, *kink-kinked*, *link-linked*, *pink-pinked*, *prink-prinked* and *wink-winked*.

inconsistency of low- than high-frequency verbs on the *token* measure. Such an account illustrates the impact of graded effects of the consistency of the present-to-past-tense transformation central to single-mechanism accounts, comparable to the effects of consistency in the transformation from orthographic body to phonological rime observed in studies of reading aloud (Jared, 1997; Jared, 2002).

Probably the most surprising finding that emerged from the present experiment was the lack of significant effects of lemma frequency upon either RTs or error rates in the Picture Inflection task. It should be noted, however, that the IPNP action-picture naming RTs for the present stimuli did not reveal a significant frequency effect either. Indeed, in a series of regression analyses of the full set of IPNP pictures, Szekely et al. (2005) observed a small but significant positive relationship between frequency and action-picture naming latency, although the same study yielded a strong negative relationship between frequency and RTs for naming *object* pictures. Following Szekely et al. (2005), we attribute the attenuated frequency effects observed in action relative to object naming to a preponderance of light verbs in the higher-frequency ranges. Light verbs are those that appear in many diverse linguistic contexts, and their high frequency tends to reflect the sum of all their different senses. For example, the verb "give" can have many different meanings according to the context in which it occurs, whereas the semantic representation of the noun "tree" is relatively constant.

In terms of the current past-tense generation tasks, the presentation of an isolated verb stem activates a form-based representation that presumably inherits the frequency of all senses of that item. As such, the overall frequency for a light verb will correspond to its level of activation in much the same way as for a verb with a more specific meaning. In contrast, in the Picture Inflection task, when a particular meaning of a verb is instantiated in an action picture, this will most likely represent only a subset of the meanings associated with that item in the case of light verbs (e.g., "to give a present", as depicted in the present study, versus "to give up" or "to give in" or "to give a piece of one's mind" or "My dear, I don't give a damn!" etc.). As light verbs tend to be more common, this would attenuate any processing benefits otherwise obtained for the high-frequency verbs used here when presented as pictures rather than stems, eliminating the frequency effect for the Picture Inflection task. In support of this account, the number of senses listed for each verb used in the present task in the Wordsmyth Online Dictionary (http://www.wordsmyth.net/) tended to be greater for the high- than low-frequency items (9.1 vs 7.8 senses; $t(29) = 1.45$, $p = .08$, one-tailed). Of course, frequency effects, whether arising from lexical or semantic sources, are central to both dual- and single-mechanism accounts of inflectional morphology. Yet the light-verb explanation just offered for the absence of frequency effects in the Picture Inflection task seems more compatible with the single-mechanism account, as inflection can proceed directly from a particular instantiation of a verb's meaning. An assumption of mandatory stem retrieval in the dual-mechanism approach entails that any instantiation of a verb's meaning should inherit the full frequency associated with the context-independent lexical representation of its stem.

Experiment 1 has demonstrated that the significant disadvantage for irregular verbs reliably observed in the standard Stem Inflection task can entirely disappear when the task becomes one of producing the past tense from meaning. This pattern of results favours single- over dual- mechanism models of inflectional morphology. There were, however, a number of other differences between the characteristics of performance in these two tasks that might qualify this conclusion. The first is that RTs were substantially slower in Picture Inflection than Stem Inflection. Although this difference is not unexpected in light of the substantial difference between pictures and words in terms of visual complexity and recognition procedures, it is possible that the regularity effect of interest is simply attenuated at longer latencies. A second difference concerns the relative incidence of "other name" errors. Such errors were far more prevalent in Picture Inflection than Stem Inflection, inevitably so given the inherently greater ambiguity associated with assigning a specific verb to an action picture relative to simply reading a verb stem. Of more concern, however, is the fact that other names were more often assigned to pictures of regular- than irregular-verb referents. This suggests that, in spite of the fact that the regular and irregular items did not differ significantly in IPNP name agreement, it was nevertheless more difficult to establish the specific intended referent action for the former than the latter. This aspect of the stimuli might have counteracted the emergence of a regularity effect in correct RTs for Picture Inflection. To overcome these limitations, a familiarisation procedure was used in Experiment 2.

Experiment 2

The purpose of Experiment 2 was to rule out alternative explanations for the absence of a regularity effect in the Picture Inflection task in Experiment 1, in terms of either long RTs or greater ambiguity in the picture-verb associations for regular than irregular items. The manipulation employed in our attempt to achieve these goals was a picture-name familiarisation task, in which participants were initially exposed to the action pictures with their intended verb-stem labels, which they read aloud at their leisure. This procedure was intended to decrease both response times and the number of "other name" responses in the Picture Inflection task. As there is some evidence that prior production of phonological components of a verb's stem may prime a regularisation response (Stemberger, 2004), a comparable verb-name familiarisation task preceded the Stem Inflection task in order to equate the potential for any such phonological priming across the two stimulus conditions.

Method

Participants

A further 24 members of the Cognition and Brain Sciences Unit volunteer panel were paid for their participation in this study. As in Experiment 1, all were native speakers of British English, aged between 18 and 40 years.

Stimuli

The stimuli for Experiment 2 were identical to those used in Experiment 1.

Procedure

Equipment, error recording and condition assignment were as per Experiment 1. Immediately before the Stem Inflection task, participants were familiarised with each verb stem, with no mention made of the subsequent past-tense task. Each familiarisation trial began with a white cross on a black screen for 500 ms, followed by central presentation of the verb stem. The participants' task was simply to read aloud the verb stem with no time pressure. The verb stem remained on the screen until response or for a maximum of 4000 ms. Similarly, immediately before the Picture Inflection task, participants were familiarised with each picture name. Participants were instructed that they would need to know the name of each action picture in order to complete a subsequent task, however no mention was made of the requirement to generate the past tense. Each familiarisation trial began with a white cross on a black screen for 500 ms, followed by central presentation of the line drawing with the verb stem printed underneath each picture. The participants' task was to inspect the picture and then read aloud its name, again with no time pressure. The picture and stem remained on the screen until response or for a maximum of 4000 ms. For both the verb- and picture-name familiarisation tasks, each vocal response was recorded from the onset of the trial for a period of 2000 ms after the voice-key trigger, after which a new trial then began after an interval of 420 ms. For the subsequent inflection tasks, stimulus presentation and response recording were identical to Experiment 1.

Results

Data treatment and error classification were as per Experiment 1. Item RT and error rate data for all conditions of the inflection tasks appear online as Supplementary material, with a breakdown of the different error types provided in Appendix C.

Stem Inflection

Mean correct RTs and percentage of errors of interest (regularisations plus LARC errors) Figs. 1B and 2B, respectively, with exact values provided online as Supplementary material. These data were analysed with 2 (frequency) by 2 (regularity) within-participants and between-items ANOVAs. High-frequency verbs were inflected significantly more rapidly but not reliably more accurately than low-frequency verbs, $F_1(1,11) = 13.34$, $p < .01$; $F_2(1,56) = 5.52$, $p = .02$; $\min F'(1,60) = 3.90$, $p = .05$; $F_1(1,11) = 6.71$, $p = .03$; $F_2(1,56) = 2.12$, *ns*; $\min F'(1,64) = 1.61$, *ns*. Regular verbs were inflected significantly more rapidly and more accurately than irregular verbs, $F_1(1,11) = 12.02$, $p = .01$; $F_2(1,56) = 18.02$, $p < .01$; $\min F'(1,28) = 7.21$, $p = .01$; $F_1(1,11) = 28.60$, $p < .01$; $F_2(1,56) = 14.30$, $p < .01$; $\min F'(1,55) = 9.53$, $p < .01$. The interaction between frequency and regularity was not reliable in either RTs or error rates, *F*s < 1; $F_1(1,11) = 6.71$, $p = .03$; $F_2(1,56) = 2.12$, *ns*; $\min F'(1,64) = 1.61$, *ns*.

Picture Inflection

Mean correct RTs and percentage of errors of interest (regularisations plus LARC errors) are presented in Figs. 1B and 2B, respectively, with exact values provided online as Supplementary material. These data were analysed in the same way as the Stem Inflection results. High-frequency verbs were inflected significantly more slowly than low-frequency verbs, $F_1(1,11) = 6.99$, $p = .02$; $F_2(1,56) = 6.57$, $p = .01$; $\min F'(1,38) = 3.39$, $p = .07$, but also significantly more accurately, albeit marginally so by items, $F_1(1,11) = 4.71$, $p = .05$; $F_2(1,56) = 3.27$, $p = .08$; $\min F'(1,47) = 1.93$, *ns*. Performance on irregular verbs was not reliably slower or less accurate than on regular verbs, $F_1(1,11) = 5.61$, $p = .04$; $F_2(1,56) = 1.65$, *ns*; $\min F'(1,65) = 1.38$, *ns*; $F_1(1,11) = 3.14$, *ns*; $F_2(1,56) = 1.46$, *ns*; $\min F'(1,57) = 1.00$, *ns*, irrespective of their frequency, *F*s < 1; $F_1(1,11) = 3.14$, *ns*; $F_2(1,56) = 1.46$, *ns*; $\min F'(1,57) = 1.00$, *ns*.

Stem versus Picture Inflection

To provide an assessment of task differences, these data were analysed with 2 (frequency) by 2 (regularity) by 2 (stimulus type) ANOVAs. As before, regularity and frequency were treated as within-participants and between-items factors, while stimulus type was treated as a between-participants and within-items factor. With respect to RTs, Stem Inflection did not differ reliably from inflection from Picture Inflection, $F_1(1,22) = 1.66$, *ns*; $F_2(1,56) = 134.12$, $p < .01$; $\min F'(1,23) = 1.64$, *ns*. The opposing directions of the frequency effect for Stem Inflection compared to Picture Inflection resulted in a significant interaction between stimulus type and frequency, $F_1(1,22) = 17.98$, $p < .01$; $F_2(1,56) = 14.86$, $p < .01$; $\min F'(1,68) = 8.14$, $p = .01$. Most importantly, the reduction in the regularity effect for Picture Inflection relative to Stem Inflection was highly significant, $F_1(1,22) = 17.17$, $p < .01$; $F_2(1,56) = 19.86$, $p < .01$; $\min F'(1,59) = 9.21$, $p < .01$. Given the lack of a significant frequency-by-regularity interaction in RTs for both stimulus conditions, it is not surprising that the three-way interaction with stimulus type did not approach significance, *F*s < 1.

Turning to errors, there were significantly more regularisation plus LARC errors in Stem Inflection than Picture Inflection, $F_1(1,22) = 12.79$, $p < .01$; $F_2(1,56) = 8.81$, $p < .01$; $\min F'(1,72) = 5.22$, $p = .03$. There was no reliable frequency effect on these error types, *F*s < 1. As observed in RTs, the reduction in the regularity effect seen for Picture Inflection relative to Stem Inflection was highly significant, $F_1(1,22) = 16.85$, $p < .01$; $F_2(1,56) = 10.65$, $p < .01$; $\min F'(1,37) = 6.53$, $p = .01$. The three-way interaction between regularity, frequency and stimulus type did not approach significance, $F_1(1,22) = 1.80$, *ns*; $F_2 < 1$.

A consideration of the "other name" error rates observed in Experiment 2, presented in Fig. 3B, indicates that such responses were no longer reliably more common for Picture Inflection than for Stem Inflection, $F_1(1,22) = 2.22$, *ns*; $F_2(1,56) = 6.87$, $p = .01$; $\min F'(1,37) = 1.68$, *ns*. The incidence of such responses was not reliably affected by either frequency, $F_1(1,22) = 4.10$, $p = .06$; $F_2(1,56) = 1.42$, *ns*; $\min F'(1,78) = 1.05$, *ns*, or regularity, *F*s < 1, nor did the three-way interaction with stimulus type approach significance, *F*s < 1.

Impact of familiarisation on inflection

In order assess the effect of familiarisation on inflection from form versus meaning, data from Experiments 1 and 2 were analysed together with familiarisation included as a between-participants and within-items factor. The analysis revealed that a significant overall decrease in RTs associated with familiarisation, albeit marginally so across participants, $F_1(1,44) = 3.08$, $p = .09$; $F_2(1,56) = 100.51$, $p < .01$; $\min F'(1,47) = 2.99$, $p = .09$. This decrease was qualified by a significant interaction with stimulus type, $F_1(1,44) = 9.54$, $p < .01$; $F_2(1,56) = 100.81$, $p < .01$; $\min F'(1,52) = 8.72$, $p = .01$. The basis for this interaction can be seen in a comparison of Figs. 1A and 1B: familiarisation caused a small increase in RTs for Stem Inflection, $F_1 < 1$; $F_2(1,56) = 22.25$, $p < .01$; $\min F' < 1$, but a substantial decrease in RTs for Picture Inflection, $F_1(1,44) = 11.72$, $p < .01$; $F_2(1,56) = 64.45$, $p < .01$; $\min F'(1,60) = 9.92$, $p < .01$. No other effects involving familiarisation were reliable.

In terms of regularisation plus LARC error rates, familiarisation caused a reliable overall increase in such errors, $F_1(1,44) = 4.91$, $p = .03$; $F_2(1,56) = 12.50$, $p < .01$; $\min F'(1,76) = 3.53$, $p = .06$, and, as would be expected, this effect was most apparent for irregular items, $F_1(1,44) = 6.26$, $p = .02$; $F_2(1,56) = 12.50$, $p < .01$; $\min F'(1,83) = 4.17$, $p = .04$. As can be seen in Fig. 2A and 2B, a reliable three-way interaction with stimulus type, $F_1(1,44) = 4.79$, $p = .03$; $F_2(1,56) = 4.48$, $p = .04$; $\min F'(1,99) = 2.31$, *ns*, indicated that the increased error rate for irregular items produced by familiarisation was more marked for Stem Inflection, $F_1(1,44) = 8.80$, $p = .01$; $F_2(1,56) = 7.09$, $p = .01$; $\min F'(1,100) = 3.93$, $p = .05$, than Picture Inflection $F_1 < 1$; $F_2(1,56) = 9.93$, $p < .01$, $\min F' < 1$. No other effects involving familiarisation were reliable.

With respect to the incidence of "other name" errors, their significant reduction due to familiarisation, $F_1(1,44) = 49.95$, $p < .01$; $F_2(1,56) = 35.53$, $p < .01$; $\min F'(1,100) = 20.76$, $p < .01$, was more apparent for Picture Inflection than Stem Inflection, $F_1(1,44) = 54.82$, $p < .01$; $F_2(1,56) = 32.70$, $p < .01$; $\min F'(1,98) = 20.48$, $p < .01$, and for regular than irregular items, $F_1(1,44) = 11.43$, $p < .01$; $F_2(1,56) = 3.81$, $p = .06$; $\min F'(1,87) = 2.86$, $p = .10$. As is apparent in Fig. 3A and 3B, a reliable three-way interaction between these factors, $F_1(1,44) = 9.15$ $p < .01$; $F_2(1,56) = 4.88$, $p = .03$; $\min F'(1,97) = 3.18$, $p = .08$, reflects the fact that the reduction in "other name" errors after familiarisation in the Picture Inflection task was more pronounced for regular items, $F_1(1,44) = 57.25$, $p < .01$; $F_2(1,56) = 20.90$, $p < .01$; $\min F'(1,89) = 15.31$, $p < .01$, than irregular items, $F_1(1,44) = 9.95$, $p = .01$; $F_2(1,56) = 10.44$, $p < .01$; $\min F'(1,98) = 5.09$, $p = .03$. No other effects involving familiarisation were reliable.

Discussion

The results of Experiment 2 validate those of Experiment 1 in that once again, for both response times and regularisation/LARC error rates, the significant regularity effects observed in standard Stem Inflection were eliminated in Picture Inflection. The inclusion of a stem/picture-name familiarisation procedure achieved its intended goals, in that Stem and Picture Inflection were now comparable in terms of overall RTs, and the "other name" errors in Picture Inflection were not only substantially reduced, but were evenly distributed amongst regular and irregular items. Hence the results of Experiment 2 demonstrate that the absence of a regularity effect in Picture Inflection in Experiment 1 cannot be attributed to long RTs or to greater referent ambiguity for regular than irregular items.

Although both experiments yielded a striking reduction in the regularity effect when people produced the past tense from meaning rather than form, the use of the familiarisation procedure in Experiment 2 did produce a few noteworthy changes relative to the results observed in Experiment 1. One consequence of the familiarisation procedure was that the null effect of frequency for Picture Inflection in the RTs of Experiment 1 actually became a significant reversal in Experiment 2, resulting in a reliable interaction between frequency and stimulus type. Although this may be at least partially due to a speed accuracy trade-off, as a standard frequency effect was observed in LARC/regularisation error rates for Picture Inflection, the significantly reversed frequency effect seen in Picture Inflection RTs parallels the effect observed for the full set of action pictures in the IPNP corpus reported by Szekely et al. (2005). As before, we attribute this result to the preponderance of light verbs in the higher-frequency set. We presume that picture familiarisation produced a reverse frequency effect because prior instantiation of a particular sense of a light verb's meaning (e.g., "to give a present") functioned more effectively to constrain the meaning evoked upon the second presentation of the picture (during the Picture Inflection task), thereby enhancing the light-verb phenomenon responsible for the attenuation of the frequency effect during Picture Inflection in Experiment 1.

Another consequence of the familiarisation procedure was that prior presentation and production of the verbs' stems significantly increased the regularity effect observed in regularisation/LARC error rates in the Stem Inflection task, eliminating the regularity-by-frequency interaction observed for this measure in Experiment 1. This result is in line with Stemberger's (2004) recent observation that presentation and production of a noun containing the vowel of a verb's stem immediately preceding Stem Inflection (e.g., "The cream..." *freeze* → "froze" or "freezed") significantly increased the likelihood of a regularisation error relative to prior presentation and production of a phonologically unrelated word (e.g., "The slot..." *freeze* → "froze" or "freezed"). What is particularly interesting in the present study is that prior presentation and production of the verbs' stems did not significantly increase the regularity effect in either RT or error rate in the Picture Inflection task. This suggests that phonological priming of regularisation responses from familiarisation is confined to the case where this earlier activation summates with activation from encountering the verb's stem immediately prior to inflection.

Simulation

The contrasting pattern of performance observed here for Stem versus Picture Inflection suggests that the processes involved in producing the past tense differ depending on the task used to elicit these forms. Such a result appears to be incompatible with a dual-mechanism perspective if this account is assumed to posit mandatory activation of a verb's stem form before a speaker can produce its past tense. Although we have suggested that inflection via a single mechanism should produce such differences in the size of the regularity effect according to the nature of the input, this is yet to be demonstrated. It is certainly true that production of the past tense from meaning within a single mechanism model incorporating semantic representations does not *necessitate* intermediate retrieval of the stem; it remains possible, however, that there may still be some activation of the stem's phonological form, given the cascaded and interactive nature of processing within connectionist models. The only way to establish that the single-mechanism view can account for our results is direct simulation.

The simulation component of the study addressed this issue using a single-mechanism architecture originally reported in Joanisse and Seidenberg (1999), which implements past-tense processing in terms of the mappings between verbs' phonological and semantic forms. A key feature of this model is that it acquires a corpus of present- and past-tense verbs via four different training tasks: 'hearing', where a phonological form is the input and the model activates the corresponding semantics as output; 'speaking', in which a word's semantic form is the input and the model produces the phonological form as output; 'repeating', in which a phonological input is replicated on a phonological output layer; and 'generating', in which the present-tense form is provided as an input and the model must produce the past tense as output.

The structure of this network respects some of the key distinctions between the Stem and Picture Inflection tasks examined in the present work. For both tasks, the instruction to produce a past-tense output can be simulated by activating the 'past-tense' unit on the semantics layer. Stem Inflection maps directly onto the 'generating' task, which involves taking the phonological form of a present-tense stem (as derived from its written form) and producing its past tense. Picture Inflection is straightforwardly simulated by the 'speaking' task, on the assumption that an action picture will activate semantic features corresponding to a verb, which then trigger a corresponding phonological output. Thus, although the model was not originally conceived as a way to account for the behavioural data presented here, its architecture lends itself naturally to both task types. To examine whether these distinctions have consequences for past-tense production in the model similar to those observed in Experiments 1 and 2, we assessed the model's performance on the high- and low-frequency regular and irregular items used in these experiments. We hypothesized that differences in how the model implements Stem and Picture Inflection would result in task differences with respect to the size of the regularity effects observed.

Method

The network architecture resembles the one reported previously by Joanisse and Seidenberg (1999), with three key modifications. First, the model learned a larger corpus of words that included all items in the test set. This corpus was based on UK speech, to correspond better to the participants in Experiments 1 and 2 (the original model learned US pronunciations). Second, unit activation was implemented in a slightly different way, which better captured the dynamical nature of word processing. Third, distributed semantic representations were used. In the original simulation, a localist semantic coding scheme uniquely assigned individual units to each verb concept. This original scheme was used primarily in the interest of computational efficiency, since localist coding permitted faster training. It had also been assumed that semantic similarity amongst verbs was not a major determinant of past-tense processing, although newer evidence suggests that this dimension is indeed relevant under some circumstances (Baayen & Moscoso del Prado Martín, 2005; Gordon & Miozzo, 2008; Ramscar, 2002).

Model architecture

The structure of the model is depicted in Fig. 4. Artificial neurons ('units') were grouped into layers, which were exhaustively connected to all neurons in adjacent layers using weighted, directional connections (illustrated as arrows). Input and output layers encoded information about either the phonological or semantic form of words in the training corpus. The Phonological Input layer consisted of 180 units. It represented the phonological form of words as input to the network units using a slot-based coding scheme [CCCVVCCC-VC], where each consonant or vowel was coded as a vector of 18 phonological features that captured all relevant natural classes of phonemes in English (for further details see Joanisse & Seidenberg, 1999). The Phonological Output layer used an identical coding scheme.

The Semantic layer consisted of 251 units, where each word's semantic form was encoded as a unique pattern of activation across these units. A random but structured coding scheme was used, adapted from Plaut (1997). First, a prototype pattern was created by randomly setting 20 of a possible 250 units to a value of 1 ('on'). Next, 18 exemplar patterns were created by randomly setting the 'on' units in the prototype to a value of 0 ('off') at a probability of .50. Each exemplar was also required to differ from all other exemplars by a minimum of 3 units. This process was repeated for 120 different random prototypes, generating a total of 3000 semantic representations that varied in their similarity to one another. Each present-tense form was as-

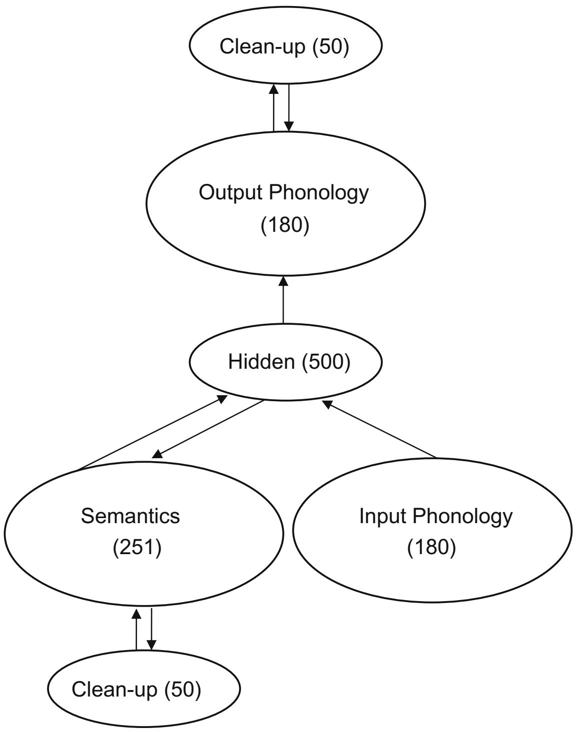

Fig. 4. Structure of the modified version of the Joanisse and Seidenberg (1999) model used in the present simulations. The number of units in each layer is provided in parentheses. Arrows indicate full interconnectivity in the indicated directions.

signed at random to an exemplar, which was used to represent it. An additional unit in the semantics layer was also used to represent the concept of 'past', such that this unit was set to 'on' for a word's past-tense form, and 'off' for a word's present-tense form.

The semantic coding scheme described here was not intended to simulate fine-grained aspects of semantics, such as imageability, animacy or density. Rather, it captured in a very broad way the idea that multiple semantic features are used to encode our knowledge of any given concept, and that these features participate in representing other concepts as well (Cree, McRae, & McNorgan, 1999; Harm & Seidenberg, 2004; Rogers & McClelland, 2004). Critically, no single unit uniquely corresponded to any single word form, and as such there is no sense in which this coding scheme has implemented a 'lexicon' in the classic, symbolic processing sense.

A 500 unit hidden layer was used, which mediated the connections between the semantic and phonological layers of the network. Cleanup layers (50 units) were recurrently connected to the Phonological Output and Semantic layers. These permitted the network to develop attractor-type representations in which groups of units converge on a stable output representation over the course of consecutive cycles in a given trial (Plaut et al., 1996).

Unit activation in this model was implemented using a continuous-time function, in which activity is allowed to ramp up or down based on its current state and values that are input to it (Pearlmutter, 1995; Plaut et al., 1996). Each unit thus represented a leaky integrator in which activation changed slowly over time in response to inputs, rather than instantaneously as is typically implemented in discrete-time backpropagation models (e.g., Williams & Peng, 1990). This allowed unit activation values to settle into a stable activation pattern over successive cycles in the network. An integration constant of 0.25 was used (see Harm & Seidenberg, 2004, for a fuller explanation of how continuous-time units are implemented in this model).

Stimuli and training

Eight individual networks were trained, each initialized with a different random number seed. Although the architecture and training set were held constant across networks, each had a different set of initial weights and the order of presentation of items and trials was different.

Networks were trained on 1923 English verbs drawn from the CELEX database (Baayen et al., 1993). These items were selected on the basis that their present-tense stem forms were monosyllabic. The simple past-tense form was monosyllabic for 1562 items and disyllabic for 361 items. The present- and past-tense frequencies for these items were obtained from the form frequency listing of CELEX (EFW.CD). To obtain the present-tense frequencies, we summed across all four instances listed for a given form (e.g., *to sleep, I sleep, you sleep, they sleep*). To calculate the past-tense frequencies, we summed across four instances listed for a given form (e.g., *I slept, you slept, he/she slept, they slept*), but did not include the frequency given for the past participle (e.g., *have slept*). Phonological forms were taken from the CELEX transcriptions and reflect British English pronunciation.

Prior to training, weights were initialized by assigning random values to each (range: −0.1 to 0.1). On each trial, a task type was first selected at random but with different probabilities: 'speaking' (40% of trials), 'hearing' (40%), 'repeating' (10%) and 'generating' (10%). Next a training item was selected probabilistically based on the input form's natural log CELEX frequency. Input values were presented to the network for six cycles, during which time all other unit values were permitted to change as a function of their inputs. The network was then run for an additional six cycles, with target values set for the last two of these cycles, for a total of 12 cycles per training trial. Connection weights were adjusted by comparing output values to target values using a cross-entropy error calculation and then modifying weights based on this unit error using a variant of recurrent backpropagation through time adapted for continuous-time (Pearlmutter, 1995; Plaut et al., 1996). Learning rate was set to 0.05. Networks were trained over 4.5 million trials, the approximate point at which their error rates reached asymptote.

Results

The network's performance on the test items from Experiments 1 and 2 was examined for generating and speaking trials, which simulated Stem Inflection and Picture Inflection respectively. Performance was assessed by comparing the outputs of the Phonological Output layer to target values. Following previous models (Harm & Seidenberg, 2004; Plaut et al., 1996; Seidenberg & McClelland, 1989) sum square error (SSE) across the phonological output units at the 12th processing cycle for correctly produced items was taken as an analogue of RTs, and these values are provided for individual items for each task online as Supplementary material. Accuracy was determined phoneme-by-phoneme, using a nearest neighbour scoring methodology in which the phoneme that most closely matched the actual output pattern was selected as the winner based on a Euclidean distance function. If the network failed to generate the correct output for one or more phonemes in the target word, this was considered an error. We observed very low error rates on both tasks (3.44%, or 33 errors out of a total of 960 test trials), precluding statistical analyses of error types.

Generating trials

Mean SSE for correctly produced items in the generating trials are presented in Fig. 5. These data were analysed with 2 (frequency) by 2 (regularity) within-participants and between-items ANOVAs. The results showed significant effects of frequency, $F_1(1,7) = 15.39$, $p = .01$; $F_2(1,56) = 5.98$, $p = .02$; $\min F'(1,49) = 4.31$, $p = .04$, and regularity, $F_1(1,7) = 92.74$, $p < .01$; $F_2(1,56) = 60.28$, $p < .01$; $\min F'(1,35) = 36.53$, $p < .01$, mirroring the effects obtained in human performance for Stem Inflection. The interaction between frequency and regularity was also reliable, $F_1(1,7) = 8.85$, $p = .02$; $F_2(1,56) = 5.54$, $p = .02$; min-$F'(1,36) = 3.41$, $p = .07$, an effect that failed to materialise in the human RT data, although this interaction was observed in the error data from Experiment 1. In terms of errors, all 18 were on irregular items, with 14 of these to low-frequency verbs, and the majority involved vowel changes.

Speaking trials

Mean SSE for correctly produced items in the speaking trials are presented in Fig. 5. These data were analysed in the same manner as the generating trials. In contrast to the human data, the frequency effect was significant, with high-frequency words processed more efficiently than low-frequency words, $F_1(1,7) = 58.30$, $p < .01$; $F_2(1,56) = 24.18$, $p < .01$; $\min F'(1,47) = 17.09$, $p < .01$. In contrast to the generating trials, and comparable to human performance in Picture Inflection, the regularity effect was not significant, $F_1(1,7) = 2.37$, *ns*; $F_2(1,56) = 1.91$, *ns*; min-$F'(1,29) = 1.06$, *ns*. The interaction between frequency and regularity was also not significant, F_1 $(1,7) = 1.27$, *ns*; F_2 < 1. Overall, 15 errors were made on the speaking trials, and these were quite evenly distributed across regularity for both high-frequency (1 regular: 2 irregular) and low-frequency (5 regular: 7 irregular) items, with the majority involving vowel and/or consonant changes.

Generating versus speaking trials

Mean SSE for correctly produced items in the generating and speaking trials can be compared in Fig. 5. For an assessment of task differences, the data were analysed with 2 (frequency) by 2 (regularity) by 2 (stimulus type) ANOVAs. As before, regularity and frequency were treated as within-participants and between-items factors, while stimulus type was treated as a within-participants and within-items factor (given that the same network was assessed on each task). Overall SSE did not differ reliably across the generating and speaking trials, $F_1(1,7) = 3.36$, *ns*; $F_2(1,56) = 3.02$, $p = .09$; $\min F'(1,27) = 1.59$, *ns*. There was a significant interaction between task type and frequency, $F_1(1,7) = 14.89$, $p = .01$; $F_2(1,56) = 7.65$, $p = .01$; $\min F'(1,41) = 5.05$, $p = .03$, denoting a stronger effect in the speaking than generating trials. Most importantly, the reduction in the regularity effect for speaking relative to generating trials was highly significant, $F_1(1,7) = 14.41$, $p = .01$; $F_2(1,56) = 38.48$, $p < .01$; $\min F'(1,13) = 10.48$, $p = .01$, as seen in the human data of Experiments 1 and 2. Although a significant frequency-by-regularity interaction was confined to the generating trials, the three-way interaction with task did not approach significance, $F_1(1,7) < 1$; $F_2(1,56) = 1.85$, *ns*.

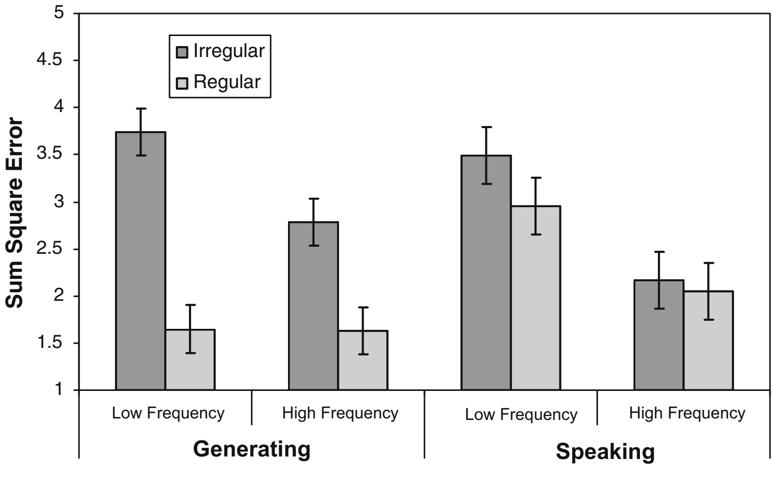

Fig. 5. Mean square error for correct trials in the modified version of the Joanisse and Seidenberg (1999) model as a function of frequency and regularity for "generating" trials (simulating Stem Inflection) and "speaking" trials (simulating Picture Inflection). Error bars represent ±95% within-participant confidence intervals for each task (Masson & Loftus, 2003, Eq. (7), p. 212).

Discussion

These simulation results suggest that the Joanisse and Seidenberg (1999) single-mechanism model provides a good fit to the human data we have reported. When the model was required to generate the past tense from form in the generation trials (analogous to our Stem Inflection task), a highly significant effect of regularity was obtained. This result is similar to that obtained by Hoeffner (1992; Hoeffner & McClelland, 1993) in a simulation of Stem Inflection in a single-mechanism model incorporating semantics in which input and output phonology were not separated. The critical result from the present simulation is that no reliable effects of regularity emerged when the model was required to generate the past tense from meaning (analogous to our Picture Inflection task). The significant reduction in the magnitude of the regularity effect for speaking trials relative to generating trials therefore precisely mirrors the contrast between performance in the Picture and Stem Inflection tasks.

There were nevertheless two points of divergence between the simulation results and the behavioural data that need to be considered. Although the reliable interaction between frequency and regularity seen in generating trials was not apparent in the latency data for Stem Inflection from Experiment 1, it was present in the error data, and hence it would seem that this effect is merely reflected in different dependent variables in the model versus human speakers. The contrast between the robust frequency effect in the model's speaking trials and the absence of this effect in the Picture Inflection data of Experiment 1 is likely to reflect the fact that all senses of a verb were activated as input for the simulation, whereas, as outlined earlier in our consideration of light verbs, human speakers were presented with a picture that instantiated only one of a number of possible senses for a given verb; this would attenuate the frequency effect given that higher-frequency verbs tended to have more alternative senses.

Overall, with respect to the critical task difference in the magnitude of the regularity effect that we obtained in human speakers, performance of this computational model clearly confirms our expectations regarding the extent to which activation of stem phonology is required by the task. In generating trials, the phonological form of the stem is instantiated across the input phonological nodes of the network. Exposure to the task of simple repetition during training predisposes the model to reproduce this stem across the output phonological nodes. The preponderance of regular past-tense forms in the model's training corpus means that reproduction of the stem (plus the addition of one of the allophones that instantiate –ed) is by far the most common transformation in generation trials. Hence, for the model, when presented with the stem form of a verb as phonological input, the strategy of reproducing that form (plus –ed) as output is the most efficient approach, and correct except for irregular verbs that require some adjustment to the stem (eg., *take-took*) or suppression of the –ed allophone (e.g., *hit-hit*). This efficiency advantage for regular verbs produces a substantial regularity effect.

The model's performance on speaking trials is quite different. Here the network must map a distributed semantic pattern that includes the concept of "pastness" onto its corresponding phonological output form, and in this case, there is no reliable effect of regularity. This difference is particularly noteworthy because the cascaded and interactive processing inherent in connectionist models of this type might well result in some activation from semantics to the phonological form of the verb stem. If this did occur, it was not sufficient to hamper the processing of irregular items. Similarly, although we might have expected the frequent co-occurrence between the semantic representation of "pastness" and the allophones of –ed to facilitate processing of regular items, this was not the case. Hence it would seem that, once the semantic representation of an action combines with the concept of pastness, activation of the corresponding phonological output representation takes the form of "direct access" (Hoeffner, 1992), in the sense that no intermediate retrieval of the phonological form of the stem is required, and therefore performance does not differ significantly across regular and irregular items.

The simulations that we have provided here are only possible in the context of implementations of the single-mechanism view that incorporate semantic representations (Hoeffner, 1992; Joanisse & Seidenberg, 1999), and could not have been attempted using more traditional models of present-to-past-tense mappings framed only in terms of phonology (Daugherty & Seidenberg, 1992; Daugherty & Seidenberg, 1994; Rumelhart & McClelland, 1986). This is not to say, however, that the model provides a complete account of the tasks of Stem and Picture Inflection. With respect to Stem Inflection, the computational simulation is more akin to auditory presentation of the stem, and does not encompass the additional reading processes involved in the case of visual presentation of the stem, as used here. In terms of Picture Inflection, as in all extant models of speech production (Dell et al., 1993; Levelt, Roelofs, & Meyer, 1999; Plaut & Kello, 1999), the current model does not address the processes involved in recognising the depicted action. Nonetheless, the close match between the human and simulation data that we obtained highlights the benefits of implemented computational models: these simulations not only *postulate* mechanisms by which aspects of language might be learned and processed, but *demonstrate* that such mechanisms can work.

General discussion

The aim of the present research was to question and evaluate the assumption that the processes engaged when participants are required to generate the past tense from form, as in the widely-used Stem Inflection task, are comparable to those involved in generation of the past tense from meaning, implemented here as a novel Picture Inflection task. Across two experiments, the robust regularity effect observed in Stem Inflection was significantly reduced when the task became Picture Inflection, with no reliable differences between regular and irregular items obtained. Furthermore, a new instantiation of the Joanisse and Seidenberg (1999) computational model of verb processing, with tasks analogous to Stem and Picture Inflection yielded the same pattern. This demonstrates that single-mechanism models of inflectional morphology that incorporate a semantic system are capable of accurately simulating performance in past-tense generation tasks where the input constitutes either form or meaning.

Our results agree with the previous, largely informal, observation that regularisation error rates to irregular verbs are far lower in natural language situations relative to elicitation tasks involving presentation of the verb stem. These observations also align with reports that a very recent encounter with the stem of an irregular verb in conversation tends to induce subsequent regularisation of the past tense in both children (Marcus et al., 1992) and patients with semantic dementia (Patterson et al., 2001). Our findings also converge with those of a previous study reporting that the regularity effect observed in Stem Inflection was reduced or eliminated when speakers were asked to produce the past tense to complete a meaningful sentence context (Okrent, 2004). The present results thus add to mounting evidence that past-tense generation in the standard Stem Inflection task is not a good analogue of past-tense production in natural language.

Stem priming effects in inflection

Before turning to theoretical implications of the present results, it is perhaps worth reflecting briefly upon the source of the robust regularity effect observed in Stem Inflection, as the absence of the effect in our Picture Inflection task establishes that it is not a guaranteed consequence of the regular/irregular distinction. The standard account of this effect within either dual- or single-mechanism models of inflectional morphology is that the difference results from the nature of the mappings involved, although the distinction is dichotomised in the former and graded in the latter. There is, however, another potential source contributing to the effect, which is that of simple response conflict induced by seeing or hearing a verb's stem. Given that both reading aloud (if written) and repetition (if spoken) are more natural responses to a verb stem than inflection, the language processing system will have some inclination to reproduce the stem form, which will produce response conflict and increase the likelihood of regularisation errors for irregular verbs whose past-tense forms do not incorporate the stem.

According to this view, presentation of the stem in the standard Stem Inflection task actually results in automatic priming of its phonology. Evidence supporting such an account comes from a study of Dutch inflection by Tabak, Schreuder, and Baayen (2005). In the Stem Inflection task, these authors reported that, whereas frequency of the present-tense form was inhibitory, frequency of the past-tense form was facilitatory. By contrast, in a novel task of present-tense production from simple past-tense, this pattern of frequency effects was reversed. Thus, there appears to be a general interference effect between the presented and to-be-produced stimuli in such form-to-form translation tasks. This competition between stimulus and response should have particularly deleterious consequences for irregular items in light of their generally greater divergence between the phonological forms of the stem and past tense (e.g., *think-thought*) relative to that seen for regular items (e.g., *walk-walked*).

In addition to priming from perception of the stem in standard elicitation tasks, results from the familiarised Stem Inflection task used in Experiment 2 suggest that there is also priming from prior presentation and production of a verb's stem that exacerbates the processing disadvantage seen for irregular items. In the current study, this took the form of long-lag repetition priming, such that presentation and production of irregular verb stems like *blow* during familiarisation increased the likelihood of "blowed" as a response to *blow* in the following Stem Inflection task. Stemberger (2004) has recently reported a similar effect in the form of immediate subword phonological priming, such that presentation and production of a noun like *cream* that contained the vowel of a subsequently presented irregular verb stem increased the likelihood of "freezed" as a response to *freeze* for that Stem Inflection trial. This latter finding suggests that the effect of familiarisation upon Stem Inflection observed in our Experiment 2 arises from a general phonological locus.

Moreover, it would seem that the negative consequences of automatic priming of stem phonology during the Stem Inflection task are not entirely confined to irregular verbs. Matthews and Theakston (2006) considered the performance of a large group of 3- to 9-year old children when the participants were required to inflect verbs for past-tense and nouns for plurality, with a particular focus on errors of 'zero marking' (no change to the stem). With respect to verb inflection, they found that such errors for regular verbs were more likely in cases where the stem already ended with a –d or –t (e.g., *load or twist*) than when it did not (e.g., *guess or cough*). The same pattern was apparent when the task was to generate noun plurals: zero-marking errors to regular nouns were more likely where the stem already ended with a sibilant (e.g., *horse or prince*) than when it did not (e.g., *grape or beard*). These results suggest that when a task involves presentation of the stem form for inflection, phonological activation of the stem form competes for production with that of the correct response.

Implications for models of inflectional morphology

According to our interpretation of the Words and Rules dual-mechanism theoretical framework, production of a past-tense verb requires two things: something to activate the stem in the lexicon plus an internal representation of the grammatical feature of past tense. The source of the latter is presumably always semantic and syntactic, whether it comes from the speaker's own intention to refer to something that has already happened or from a sentence context that requires the past tense or from an experimenter's instruction that this is the desired form of response. The source of lexical activation of the stem, on the other hand, can come entirely from meaning (as in spontaneous speech) or entirely from form (as in the Stem Inflection task). As far as we understand the workings of dual-mechanism models, the source of the stem activation should have no consequences for the nature of the subsequent processing, such that a regularity effect would be expected in any past-tense generation task. The results of our Picture Inflection task are clearly at odds with this expectation.

Theoretical frameworks are always evolving; and just as dual-mechanism models of inflection have been somewhat modified to account for the graded effects of consistency – by allowing sub-word activation within the lexicon (e.g., Pinker, 1999) and the storage of at least some regular past-tense forms (e.g., Ullman, 2001) – it seems likely that they could be elaborated to account for the absence of a regularity effect in inflection from meaning. For example, allowing direct access from semantics to the irregular past-tense forms stored in the lexicon might suffice, if the time required for this process plus rule blocking roughly equates to the duration of stem retrieval plus rule application for regular items. We suspect that accounting for the effects of familiarisation in the Stem Inflection task might prove somewhat more challenging for this approach. Stems are, of course, activated in the lexicon for all items in this task, and one might have thought that prior exposure to the stem via the familiarisation procedure would speed lexical access during Stem Inflection, resulting in facilitation for both regular and irregular verbs. The results, however, demonstrate that familiarisation in Stem Inflection was associated with slower responses and more errors, particularly for irregular verbs.

Turning to the single-mechanism framework, both the presence of a robust regularity effect in Stem Inflection and its attenuation in Picture Inflection fit comfortably within implementations of this theoretical approach that incorporate semantic representations (Hoeffner, 1992; Joanisse & Seidenberg, 1999), as demonstrated by the simulation data we have provided. The regularity effect observed in the model's performance on generating trials (comparable to Stem Inflection) seems to arise not only due to the relative atypicality of the present-to-past-tense mappings for irregular verbs, but also as a consequence of partial activation of the output phonology of the stem form that is presented as the input. These two factors push the process of creating a response in the direction of stem+-ed, thereby slowing responses and increasing error rates for irregular relative to regular verbs. In contrast, the absence of a reliable effect of regularity upon the model's performance on speaking trials (analogous to Picture Inflection) indicates that when the semantics of a verb combine with the concept of "pastness", then the appropriate past-tense form may be generated directly – that is, without recourse to the phonological form of the stem.

This single-mechanism approach also seems well suited to account for the effects of familiarisation that we observed in the Stem Inflection task. Prior presentation and production of the verbs' stem forms in the familiarisation task should raise their activation at the level of output phonology. This would then enhance the automatic activation of stem phonology in the subsequent Stem Inflection task, increasing the conflict between correct and regularisation responses to irregular verbs and inflating the regularity effect. Of course, the viability of this proposal would have to be assessed via direct simulation. Moreover, despite the fact that this model included only verbs in its vocabulary, its use of distributed representations at the level of output phonology means that prior production of a noun containing the vowel of an irregular verb's stem could also function to increase response conflict, and thus promote regularisation errors (Stemberger, 2004). In addition, given that the representations of this model are learnt through exposure, the phonological properties of verb and noun stems may well exert an influence on the rate of zero-marking errors observed at certain points during training (Matthews & Theakston, 2006).

Conclusion

The present experiments replicated the frequently reported disadvantage, in response times and/or accuracy, for irregular relative to regular verbs when normal speakers are presented with the verbs' stem forms and asked to generate their past-tense forms. The novel and crucial finding of this study is that it appears to be this particular version of past-tense elicitation that promotes the regularity effect. The difference between regular and irregular verbs was no longer apparent when the stimulus input was an action picture rather than the verb stem, demonstrating that at least partially different processes underlie past-tense generation from form and from meaning. The patterns of performance reported here were successfully simulated within a single-mechanism model of inflection incorporating distributed semantic representations.

More generally, our study indicates that conclusions concerning the mechanisms involved in inflectional morphology drawn from performance in standard form-based elicitation tasks do not necessarily generalise to the processes underlying past-tense generation from meaning, which seem more akin to those supporting spontaneous speech. This outcome should perhaps serve as a reminder to all language researchers to reflect, from time-to-time, on the following question: to what extent are the artificial language tasks that we use in the laboratory representative of natural language processing?

Acknowledgments

The research reported here was supported by an NIMH-IBSC Grant (MH64445). We are grateful to Matt Lambon Ralph, Jay McClelland and Mark Seidenberg for helpful discussions concerning the issues considered in this paper; and to Elisa Cooper-Pye and Lynne Hemingway for research assistance.

Appendix A

Verb stems used in the present study. Corresponding action pictures used can be downloaded from: http://crl.ucsd.edu/~aszekely/ipnp/actobj/getpics/getpics.html.

Low frequency		High frequency	
Irregular	Regular	Irregular	Regular
bite	beg	catch	climb
blow	carve	cut	cry
dig	cook	drive	fill
grind	cross	feed	look
ride	dance	fight	pick
shoot	sew	give	play
sing	shave	run	push
sink	slip	sell	reach
slide	smoke	shake	save
spread	sneeze	sit	serve
steal	stir	sleep	smile
sting	tie	stand	talk
sweep	type	teach	walk
swim	wave	throw	watch
swing	wrap	win	wish

Appendix B

Breakdown of percentage error rates according to response type for low- and high-frequency irregular and regular verbs used in the present study for both the stem and picture conditions of Experiment 1. Standard deviations are given in parentheses.

	Low frequency		High frequency	
	Irregular	Regular	Irregular	Regular
Stem				
Regularisation	3.33 (5.32)	0.00 (0.00)	1.11 (2.59)	0.00 (0.00)
LARC	2.22 (4.34)	0.56 (1.92)	0.00 (0.00)	0.00 (0.00)
Past participle	7.78 (5.57)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
No change	1.11 (2.59)	0.56 (1.92)	0.00 (0.00)	0.56 (1.92)
No response	1.11 (3.85)	0.00 (0.00)	0.56 (1.92)	0.00 (0.00)
Other name	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.56 (1.92)
Total	15.56 (9.57)	1.11 (3.85)	1.67 (4.14)	1.11 (2.59)
Picture				
Regularisation	0.56 (1.92)	0.00 (0.00)	0.56 (1.92)	0.00 (0.00)
LARC	0.56 (1.92)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)

Appendix B. (*continued*)

	Low frequency		High frequency	
	Irregular	Regular	Irregular	Regular
Past participle	10.56 (6.64)	0.56 (1.92)	0.56 (1.92)	0.00 (0.00)
No change	1.67 (3.02)	0.00 (0.00)	1.67 (4.14)	0.00 (0.00)
No response	5.00 (7.59)	2.22 (4.34)	3.33 (6.67)	2.22 (4.34)
Other name	7.78 (6.25)	16.67 (7.78)	10.00 (8.76)	21.11 (11.31)
Total	26.11 (6.64)	19.44 (8.74)	16.11 (14.34)	23.33 (11.55)

Appendix C

Breakdown of percentage error rates according to response type for low- and high-frequency irregular and regular verbs used in the present study for both the stem and picture conditions of Experiment 2. Standard deviations are given in parentheses.

	Low frequency		High frequency	
	Irregular	Regular	Irregular	Regular
Stem				
Regularisation	8.33 (7.59)	0.00 (0.00)	3.89 (5.29)	0.00 (0.00)
LARC	1.67 (3.02)	0.00 (0.00)	0.56 (1.92)	0.00 (0.00)
Past participle	12.78 (15.94)	1.11 (2.59)	3.89 (7.22)	0.00 (0.00)
No change	0.56 (1.92)	0.00 (0.00)	0.56 (1.92)	0.00 (0.00)
No response	1.11 (3.85)	0.00 (0.00)	1.67 (4.14)	0.56 (1.92)
Other name	0.56 (1.92)	0.56 (1.92)	0.56 (1.92)	0.00 (0.00)
Total	25.00 (18.45)	1.67 (3.02)	11.11 (13.73)	0.56 (1.92)
Picture				
Regularisation	1.67 (3.02)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
LARC	1.11 (2.59)	0.56 (1.92)	0.00 (0.00)	0.00 (0.00)
Past participle	3.89 (6.64)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
No change	0.00 (0.00)	1.11 (2.59)	0.56 (1.92)	0.56 (1.92)
No response	0.00 (0.00)	1.67 (4.14)	0.56 (1.92)	0.56 (1.92)
Other name	1.67 (5.77)	1.11 (2.59)	2.22 (5.92)	2.78 (3.43)
Total	8.33 (11.42)	4.44 (5.19)	3.33 (6.03)	3.89 (4.46)

Supplementary material

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jml.2009.02.002.

References

Baayen, R. H., & Moscoso del Prado Martín, F. (2005). Semantic density and past-tense formation in three Germanic languages. *Language, 81*, 666–698.

Baayen, R. H., Piepenbrock, R., & van Rijn, H. (1993). *The CELEX lexical database (CD-ROM)*. Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania.

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

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

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review, 108*, 204–256.

Cortese, M., Balota, D., Sergent-Marshall, S., Buckner, R., & Gold, B. (2006). Consistency and regularity in past-tense verb generation in healthy ageing, Alzheimer's disease, and semantic dementia. *Cognitive Neuropsychology, 23*, 856–876.

Cree, G. S., McRae, K., & McNorgan, C. (1999). An attractor model of lexical conceptual processing: Simulating semantic priming. *Cognitive Science, 23*, 371–414.

Daugherty, K., & Seidenberg, M. S. (1992). Rules or connections? The past tense revisited. In *Proceedings of the 14th annual conference of the cognitive science society* (pp. 259–264). Hillsdale, NJ: Lawrence Erlbaum.

Daugherty, K., & Seidenberg, M. S. (1994). Beyond rules and exceptions. In S. D. Lima, R. L. Corrigan, & G. K. Iverson (Eds.), *The reality of linguistic rules*. Amsterdam: John Benjamins.

Dell, G. S., Juliano, C., & Govindjee, A. (1993). Structure and content in language production: A theory of frame constraints in phonological speech errors. *Cognitive Science, 17*, 149–195.

Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments and Computers, 35*, 116–124.

Gordon, P., & Miozzo, M. (2008). Can word formation be understood or understanded by semantics alone. *Cognitive Psychology*, 30–72.

Harm, M., & Seidenberg, M. S. (2004). Computing the meanings of words in reading: Cooperative division of labor between visual and phonological processes. *Psychological Review, 111*, 662–720.

Hoeffner, J. (1992). Are rules a thing of the past? The acquisition of verbal morphology by an attractor network. In *Proceedings of the 14th annual meeting of the cognitive science society*. Hillsdale, NJ: Erlbaum.

Hoeffner, J. H., & McClelland, J. L. (1993). Can a perceptual processing deficit explain the impairment of inflectional morphology in developmental dysphasia? A computational investigation. In E. V. Clark (Ed.), *Proceedings of the 25th child language research forum*. Stanford University Press.

Jared, D. (1997). Spelling-sound consistency affects the naming of high-frequency words. *Journal of Memory and Language, 36*, 505–529.

Jared, D. (2002). Spelling-sound consistency and regularity effects in word naming. *Journal of Memory and Language, 46*, 723–750.

Joanisse, M. F., & Seidenberg, M. S. (1999). Impairments in verb morphology after brain injury: A connectionist model. *Proceedings of the National Academy of Sciences of the United States of America, 96*(13), 7592–7597.

Kuczaj, S. A. (1976). *"-ing", "-s" & "-ed": A study of the acquisition of certain verb inflections.* Unpublished Doctoral dissertation, University of Minnesota.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22*, 1–75.

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

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*, X–164.

Masson, M. E. J., & Loftus, G. R. (2003). Using confidence intervals for graphically based data interpretation. *Canadian Journal of Experimental Psychology, 57*, 203–220.

Matthews, D., & Theakston, A. L. (2006). Errors of omission in English-speaking children's production of plurals and the past tense: The effects of frequency, phonology and competition. *Cognitive Science, 30*, 1027–1052.

McClelland, J. L., & Patterson, K. (2002). Rules or connections in past-tense inflections: What does the evidence rule out? *Trends in Cognitive Sciences, 6*, 465–472.

Okrent, A. (2004). *From meaning to words: An investigation of past tense verb inflection in English comparing a form to form mapping task with a meaning to form mapping task*. Unpublished Doctoral thesis, University of Chicago.

Patterson, K., Lambon Ralph, M. A., Hodges, J. R., & McClelland, J. L. (2001). Deficits in irregular past-tense verb morphology associated with degraded semantic knowledge. *Neuropsychologia, 39*, 709–724.

Patterson, K., Lambon Ralph, M. A., Jefferies, E., Woollams, A., Jones, R., Hodges, J. R., et al (2006). "Presemantic" cognition in semantic dementia: Six deficits in search of an explanation. *Journal of Cognitive Neuroscience, 18*, 169–183.

Pearlmutter, B. A. (1995). Gradient calculations for dynamic recurrent neural networks: A survey. *IEEE Transactions on Neural Networks, 6*, 1212–1228.

Pinker, S. (1991). Rules of language. *Science, 253*, 530–535.

Pinker, S. (1994). *The language instinct*. New York, NY: William Morrow and Co.

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

Pinker, S., & Prince, A. (1988). On language and connectionism: Analysis of a parallel distributed processing model of language acquisition. *Cognition, 28*, 73–193.

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

Plaut, D. C. (1997). Structure and function in the lexical system: Insights from distributed models of word reading and lexical decision. *Language and Cognitive Processes, 12*, 767–808.

Plaut, D. C., & Kello, C. T. (1999). The emergence of phonology from the interplay of speech comprehension and production: A distributed connectionist approach. In B. MacWhinney (Ed.), *The emergence of language* (pp. 381–415). Mahwah, NJ: Erlbaum.

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review, 103*, 56–115.

Prasada, S., Pinker, S., & Snyder, W. (1990). *Some evidence that irregular forms are retrieved from memory but regular forms are rule generated*. Paper presented at the 31st annual meeting of the Psychonomic Society, Chicago.

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

Ramscar, M. J. A., & Yarlett, D. G. (2007). Linguistic self-correction in the absence of feedback: A new approach to the logical problem of language acquisition. *Cognitive Science, 31*, 927–960.

Rogers, T. T., & McClelland, J. L. (2004). *Semantic cognition: A parallel distributed processing approach*. Cambridge, MA: MIT Press.

Rumelhart, D., & McClelland, J. L. (1986). On learning the past tenses of English verbs. In J. L. McClelland & D. E. Rumelhart (Eds.). *Parallel distributed processing: Psychological and biological models* (Vol. 2, pp. 216–271). Cambridge, MA: MIT Press.

Seidenberg, M. S. (1992). Connectionism without tears. In S. Davis (Ed.), *Connectionism: Theory and practice* (pp. 84–122). London: Oxford University Press.

Seidenberg, M. S. (1993). A connectionist modelling approach to word recognition and dyslexia. *Psychological Science, 4*, 299–304.

Seidenberg, M. S., & Bruck, M. (1990). *Consistency effects in the generation of past-tense morphology*. Paper presented at the 31st annual meeting of the Psychonomic Society, Chicago.

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review, 96*, 523–568.

Stemberger, J. P. (2004). Phonological priming and irregular past. *Journal of Memory and Language, 50*, 82–95.

Szekely, A., D'Amico, S., Devescovi, A., Federmeier, K., Herron, D., Iyer, G., et al (2005). Timed action and object naming. *Cortex, 41*(1), 7–25.

Szekely, A., Jacobsen, T., D'Amico, S., Devescovi, A., Andonova, E., Herron, D., et al (2004). A new on-line resource for psycholinguistic studies. *Journal of Memory and Language, 51*, 247–250.

Tabak, W., Schreuder, R., & Baayen, R. H. (2005). The processing of regular and irregular verbs. In *Proceedings of the interdisciplinary workshop on* *the identification and representation of verb features and verb classes* (pp. 121–126). Saarbrücken, Germany.

Tyler, L. K., deMornay Davies, P., Anokhina, R., Longworth, C., Randall, B., & Marslen-Wilson, W. D. (2002a). Dissociations in processing past tense morphology: Neuropathology and behavioral studies. *Journal of Cognitive Neuroscience, 14*, 79–94.

Tyler, L. K., Randall, B., & Marslen-Wilson, W. D. (2002b). Phonology and neuropsychology of the English past tense. *Neuropsychologia, 40*, 1154–1166.

Ullman, M. T. (2001). The declarative/procedural model of lexicon and grammar. *Journal of Psycholinguistic Research, 30*, 37–69.

Ullman, M. T., Corkin, S., Coppola, M., Hickok, G., Growden, J. H., Koroshetz, W. J., et al (1997). A neural dissociation within language: Evidence that the mental dictionary is part of declarative memory, and that grammatical rules are processed by the procedural system. *Journal of Cognitive Neuroscience, 9*, 266–276.

Ullman, M. T., Pancheva, R., Love, T., Yee, E., Swinney, D., & Hickok, G. (2005). Neural correlates of lexicon and grammar: Evidence from the production, reading, and judgment of inflection in aphasia. *Brain and Language, 93*, 185–238.

Williams, R. J., & Peng, J. (1990). An efficient gradient-based algorithm for on-line training of recurrent network trajectories. *Neural Computation, 2*, 490–501.

Ziegler, J. C., Montant, M., & Jacobs, A. M. (1997). The feedback consistency effect in lexical decision and naming. *Journal of Memory and Language, 37*, 533–554.