UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Experimental Evidence Against the Dual-Route Account of Inflectional Morphology

Permalink

https://escholarship.org/uc/item/1t7743j0

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 20(0)

Authors

Hahn, Ulrike
Nakisa, Ramin C.
Bailey, Todd M.
et al.

Publication Date

1998

Peer reviewed

Experimental Evidence Against the Dual-Route Account of Inflectional Morphology

Ulrike Hahn (U.Hahn@warwick.ac.uk)

Ramin C. Nakisa (ramin@psy.ox.ac.uk)*

Todd M. Bailey (todd.bailey@psy.ox.ac.uk)*

Miranda Holmes

Denise Kemp

Laura Palmer*

Department of Psychology, University of Warwick
Coventry CV4 7AL, U.K.

*Department of Experimental Psychology
University of Oxford, South Parks Rd.
Oxford OX1 3UD, U.K.

Abstract

Inflectional morphology has figured prominently not only in debate about the nature of linguistic knowledge, but also in the foundational debate between proponents of symbolic and of connectionist accounts of cognition. We present two experiments designed to test predictions of Pinker's (1991) dual-route account of inflection, the central component of which is a symbolic rule. Contrary to the predictions of the dual-route account, we find evidence of both frequency and similarity effects on the regularization of novel items (i.e., pseudo words).

Introduction

The nature of the computational processes underlying inflectional morphology —e.g., forming the past tense or a plural— has been the subject of intense debate for over ten years. Rumelhart and McClelland's (1986) model of past tense acquisition challenged long held views about the nature of linguistic knowledge. As a connectionist model, it suggested an alternative to the notion of symbolic rules as underlying linguistic competence. Rather than generate past tense forms through the application of a rule such as V → /ed/ the network learned to associate stem and past tense forms through changes to the set of weighted connections linking stem and past tense representing units. Furthermore, this kind of pattern association governed both regular (e.g., *walk - walked*) and irregular (e.g., *sing - sang*) forms; both were produced within a single, uniform network architecture. By contrast, the rule account (in its simplest form) assumed that irregulars were stored in lexical memory as rote-learned exceptions and thus, unlike regular past tense forms, were not the result of a generative process at all.

In the ensuing debate about the adequacy of a connectionist account of past tense acquisition and production (Pinker & Prince, 1988; Pinker, 1991; MacWhinney & Leinbach, 1991; Plunkett & Marchman, 1991, 1993)[to name but a few], Pinker (1991) proposed a revised version of the rule-based account which incorporated aspects of the connectionist explanation. Specifically, Pinker's "dual-route account" of inflectional morphology retains the idea that regular past-tense production proceeds through application of a symbolic rule, and that irregular past tense forms are based on lookup in lexical memory, but the notion of lexical memory is enriched by an additional pattern associator component. This pattern associator component allows phonologically (i.e, sound-) based generalization of irregular forms to account for the empirical finding that speakers will occasionally produce novel irregular forms, such as for instance *spling - splang*, in analogy to *sing - sang, ring - rang, ...,* when presented with "nonce-" or "pseudo-" words (Bybee & Moder, 1983).

Thus, the dual-route account assumes that regular and irregular production proceeds via two separate routes or systems: irregular forms derive from lexical memory, whereas regular past tense forms are *not* stored in lexical memory, but generated by a symbolic rule. In processing, the two routes interact in the following way: given a stem to inflect, lexical memory is first searched for an irregular entry, if none is found the rule is used to generate a regular as the default. Because of the pattern associator component, an irregular response which blocks rule application can arise even if there is no specific entry for the stem in question, *if* the word is sufficiently similar to other stored irregulars to elicit an irregular response through generalization. Consequently, the account

is not only dual-route but also computationally hybrid.

It is the debate about computational architecture — symbolic vs. connectionist— which has most fuelled interest in inflectional morphology and has elevated the issue not just to a microcosm for inquiry into the nature of linguistic knowledge, but to a focal point for the challenge of Cognitive Science's founding view of cognition as symbol manipulation and more recent views of cognitive computation in connectionist terms (Smolensky, 1988; Fodor & Pylyshyn, 1988).

Over the last decade, research has put forth a variety of ways in which the dual-route account and the connectionist single-route explanation might be empirically distinguished. We summarize these in the next section, before presenting our own experiments which seek to address this issue.

Previous Work

Empirical support for single- or dual-route accounts has been sought in a number of ways. Initial computational work tried to demonstrate that adequate performance, both with respect to accuracy of learning and with respect to putative developmental patterns —i.e., so-called U-shaped learning— could be achieved by a connectionist model (Rumelhart & McClelland, 1986; MacWhinney & Leinbach, 1991; Plunkett & Marchman, 1991, 1993). Replies to this work by proponents of the dual-route account has taken the form of detailed critiques of model performance (Pinker & Prince, 1988) and of new linguistic data thought problematic or even impossible for connectionist models (Marcus, Pinker, Ullman, Rosen, & Xu, 1992; Marcus, Brinkmann, Clahsen, Wiese, Woest, & Pinker, 1995; Xu & Pinker, 1995).

Most recently, computational research has provided direct comparisons of implemented single- and dual-route models, results of which for a variety of languages favored single-over dual-route models (Nakisa & Hahn, 1996; Hahn, Nakisa, & Plunkett, 1997; Nakisa, Plunkett, & Hahn, 1998). This work has also challenged the validity of some of the intuitive arguments from linguistic data provided by the dual-route account (Hahn & Nakisa,).

The other main line of empirical work has sought to establish difference between regulars and irregulars, which are to be reflective of the two routes on which they depend. Very recently this has involved imaging studies which seek correlates of regular and irregular production directly at the level of brain function (Jaeger, Lockwood, Kemmerer, Valin, Murphy, & Khalak, 1996; Penke, Weyerts, Gross, Zander, Münte, & Clahsen, 1997; Seidenberg & Hoeffner, 1998).

However, most of the research seeking to establish differences between regulars and irregulars has focused on *lexical variables* such as frequency effects (Clahsen, Eisenbeiss, & Sonnenstuhl-Henning, 1997; Ullman, 1993; Daugherty & Seidenberg, 1992). Because the dual-route account assumes lexical storage only for irregular, but not for regular inflected forms, *only* irregular past tense forms should show frequency effects on the time necessary to produce them, in lexical decision tasks or in ratings of their "naturalness". This pattern of results has widely been obtained, though apparently not without exception ((Sereno & Jongman, 1992) cited in Marcus et al. 1995).

Generalization tasks have also been used. The dual-route account invokes similarity-based generalization (via the pattern associator component) only for irregulars, whereas application of the symbolic rule governing regulars is blind to the particular sound of a word. Thus single- and dual-route predictions for suitably chosen novel words differ. Specifically, irregular generalization, i.e., production of irregular forms for novel items, should show clear similarity effects, given that generalization is based on similar sounding irregulars in the lexicon. By contrast, regular generalization should lack similarity effects. First, similar sounding regular forms should have no influence on generalization because they are not stored in the lexicon, and thus cannot lead to the lexically driven generalization via the pattern associator component that characterizes irregular generalization. Second, because the symbolic rule is blind to word sounds, similar sounding words can have no influence on the rule-route either. Consequently, regular generalization should be independent of the phonological neighborhood of an item (i.e., those words similar in sound) and should be possible even for nonce words which bear no similarity to known words such as "ploamph". Here, conflicting results have been obtained. The first study of this kind, by Prasada and Pinker (1993), found significant effects of similarity only for irregulars but not for regulars, in accordance with the dual-route model. A further study by Lee (1995) , however, found evidence of similarity-based generalization both for irregulars *and* regulars, so contradicting the dual-route predictions.

In summary, the weight of the experimental evidence with respect to both frequency and similarity effects seems to favor the dual-route account, though results have not been entirely unequivocal. This is somewhat in contrast with the results of our own computational investigations of the dual-route account (Nakisa & Hahn, 1996; Hahn et al., 1997; Nakisa et al., 1998) which have consistently failed to provide evidence in favor of the dual-route account. Thus we sought to address these issues with two experimental studies of our own.

Experiment 1

The question addressed in this experiment was whether or not generalization of *regulars* shows similarity effects. Specifically, does generalization decrease with increasing distance of the phonologically closest known regular. The dual-route prediction is that there should be similarity effects only for irregularization, but not for regularization, because regulars are generated independently of the lexicon via the symbolic route (Prasada & Pinker, 1993).

Participants

Participants were 32 Warwick University undergraduates of whom 23 were female. One participant was dropped from the study after reporting influences of knowledge of other languages.

Materials

Nonce words were automatically generated based on phoneme transition probabilities for verb stems in English. For the entire set of nonce words the phonological distance to the 15 nearest verbs in the English lexicon was established. Phonological distance was measured as Euclidean distance between the words represented as vectors of phonological features. For the test set, only items for which *all* of these 15 nearest neighbors were *regular* were selected. Furthermore, test items were selected to form 3 categories, *'near'*, *'intermediate'*, and *'far'*, depending on the distance of the single nearest neighbor. Phonotactic 'goodness' (phoneme transition probability) was matched in spread across all three categories.

The resultant test set comprised 72 nonce words with 24 items at each distance.

Procedure

The nonce words were presented in a written elicitation task. They were embedded in short story passages which had an underlined blank space at an appropriate point of the text. Participants were asked to fill in this blank with whatever inflected form of the nonce word, supplied at the beginning of the text, they felt was appropriate. The entire questionnaire consisted of eight separate passages, each accompanied by a list of nine nonce words (three from each class assigned randomly). Each nonce word was accompanied by a made up definition which fitted the text. Only regular verbs were used within the definition. The only irregular past tense used in the story passages was "was". There were eight different versions of the questionnaire to ensure that responses were not determined by the order of the nonce-words, their 'definitions' or the story passages: each list of nonce-words was presented with four different passages and definitions and each list had two different random orders. Apart from being listed in the margin of the text, the words were also presented aurally via audio-cassette.

Results

Contrary to the predictions of the dual-route account, there were significantly fewer regular responses for items of the *far* group than of the *near* group according to a Wilcoxon signed ranks test, $Z = 3.882; p < .001$. Also significant were the differences between near and intermediate group, $Z = 3.163; p < .05.$, and intermediate and far, $Z = 2.622 p < .05$.

That is, nonce words for which the nearest (regular) neighbor was close, elicited more regular responses that words for which this neighbor was distant.

The effect of distance was carried primarily by an increase in "no-change" forms (analogous to *hit* → *hit*, as opposed to, e.g., *grit* → *gritted*). The proportion of no-change forms increased from 13.7% to 25.1% from the *near* to the *far* group, $Z = 2.948, p < .05$, Wilcoxon's signed ranks.

However, an items analysis revealed that this effect was carried by (and significant only for) those nonce words ending in dental stops (/t/ and /d/). There were no significant effects

once these items were removed, though results approached significance.

Discussion

What does this result imply? First, even in its restriction to dental stops it contradicts the dual-route account. As outlined above, the rule-route is blind to phonology. Thus words ending in dental stops should be treated no different than any other sequence of sounds. Consequently, there should be no effect of distance to the nearest (regular) neighbor.

No-change forms, the response that primarily drives the effect in our data, are strange in that they might seem to indicate not a particular irregular response (as found, for instance, with *hit* → *hit*) but a failure to inflect. However, here they cannot plausibly be put down to a simple failure to understand the task, because this provides no rationale for why they occur more frequently as a function of distance to the nearest neighbor. If, however, they are viewed as a simple refusal to inflect, then this too contradicts the predictions of the account, again because all strings are equal as potential regulars.

It is worth mentioning again that the English lexicon does contain regulars ending in dental stops, such as *pit* → *pitted* or *flit* → *flitted*, though many are irregular, either involving internal vowel changes, *sit, spit*, or nochange *hit, slit, split*. It is also worth emphasising, that these results do not seem explicable as arising from analogies to known irregulars, at least at the level of stem-similarity, given that the 15 nearest neighbors of each item were regular.

The most likely explanation for the particular sensitivity with respect to dental stops seems, at least to us, to be that items ending in dental stops *already* look like past tense forms, and that the reluctance to add the regular affix is not based on confusion but on what the literature has called *product oriented considerations*. One strand of recent work on inflectional morphology (e.g., Koepcke, 1993; Bybee, 1995) has emphasised that inflection is characterized not only by analogy between *source* (i.e., stem) forms, but also influenced by considerations about "goodness" of output or product and by analogies between *product* (i.e., inflected) forms such as known past tenses. Traces of product-orientation have been identified in diachronic contexts, but have also been invoked for explanations of production (Bybee, 1995; Koepcke, 1993). We have also noted improved model-fits through the incorporation of output oriented considerations in our modelling of German inflection (Hahn and Nakisa, 1997). The data we presented here strike us as evidence of product-oriented considerations in an experimental setting.

Because they are one contributing factor, product oriented considerations can be overridden by similarities between stems, thus leading to a lesser proportion of no-change forms among the items with close regular phonological neighbors as found in our data set.

If one accepts this explanation of our results, then they simultaneously provide a second source of evidence against the dual-route account. The dual-route account is based entirely on the idea of inflectional processes operating on

stems. There is no place in the account where considerations about products (outputs) could come into play. Crucially, this stem+process orientation is shared by connectionist models in that they, too, typically model inflection as the task of associating input (stem) and output (past tense) forms.

Viewed in this light our results would seem add to an increasing body of evidence for the role of product-oriented considerations, and thus suggest a reorientation in the conception of the task which affects both dual-route account and current connectionist single-route models.

In summary, the experimental results show distance effects (albeit for dental stops only) which contradict the dual-route account, and, furthermore the particular pattern obtained here —prevalence of no-change forms— we interpret as evidence of "product-oriented" considerations, which themselves contradict the stem+process orientation of the dual-route account.

Experiment 2

The question addressed by our second experiment is whether or not regularization of nonce words is affected by the frequency of the most similar real regular item in the lexicon. Because regular production in the dual-route account is independent of the lexicon (regular forms are generated by the symbolic rule) the properties of extant regular words, such as their frequency, should be irrelevant to the production of regular forms for both words and nonce-words.

The specific way we address this question is by examining whether ratings of "naturalness" for regularized nonce words are influenced by the frequency of their most similar lexical neighbor.

Participants

The participants were 47 University of Oxford undergraduates, with approximately equal numbers of male and female students.

Materials

Nonce words were generated by changing phonological features of real monosyllabic words ('prototypes'). Words were generated from both regular and irregular prototypes. The chosen prototypes displayed a range of different frequencies as documented in the CELEX database. [1] Nonce words were formed at two levels of distance: either through change of a single phonological feature ('distance 1'), or through change of two phonological features ('distance 2'). 171 nonce words were generated in total, of which each participant saw 36. Of these 36, 9 each were words with irregular-prototype/distance 1, irregular-prototype/distance 2, regular-prototype/distance 1, and regular-prototype/distance 2. Within each of these 4 groups of 9 nonce words, 3 were generated from a high-frequency prototype, 3 from a medium-frequency prototype and 3 from a low-frequency prototype.

---

[1] CELEX can be obtained by contacting celex@mpi.nl.

Procedure

Stimuli were presented in written form, in a questionnaire which was sent out to participants. The questionnaire provided a brief sentential context for each word, along with an irregular option and a regular option for each word. For example,

- At the moment I spling
  Last year I *splung*
  Last year I *splinged*

Participants were asked to indicate the naturalness of the sound of both of the past tense forms by circling the appropriate number on a rating scale from 1 ("really bad" or "unnatural") to 7 ("sounds perfectly good and natural").

Results

The crucial question to be asked was whether the frequency of the regular prototype affected the regular rating of a nonce word. Thus we restrict our analysis here to regular prototypes and regular ratings.

Stepwise multiple regression revealed significant effects, at the level of $p < .01$, for distance, and log-frequency, as well as an interaction between distance and frequency: first, the main effect of distance, $t = 2.84, p < .0045$, and both log-frequency $t = 3.81, p < .0002$ and squared log-frequency as a non-linear term, $t = -4.43, p < .0000$ and, finally, a significant linear interaction between frequency and distance $t = -2.52, p < .0121$.

According to this interaction, if a nonce word is *very close*, i.e., one phonological feature away from the prototype (distance 1), goodness ratings increase together with prototype frequency. For nonce words that are slightly less close to their most similar neighbor, i.e., 2 phonological features from the prototype (distance 2), goodness ratings decrease as frequency increases.

Discussion

These findings, too, contradict the predictions of the dual-route account. Because regular past-tense forms are not assumed to be stored in the lexicon at all, but generated via the rule-route, frequency of the regular prototype should have no impact.

The frequency-distance interaction, too, is interesting as it provides a rationale for why significant results might be missed in studies of this kind, simply because overall effects across different phonological distances might cancel out. Furthermore, it raises interesting questions as to its cause. One possible explanation is that generalization gradients become steeper with increased frequency as illustrated in Fig. 1.

This could be a way of interpreting Bybee's (1995) suggestions that lexical representation varies as a function of token frequency. Bybee (1995) suggests that lexical entries differ in *lexical strength* which affects ease of access, ability to serve as the basis of morphological relations, and resistance
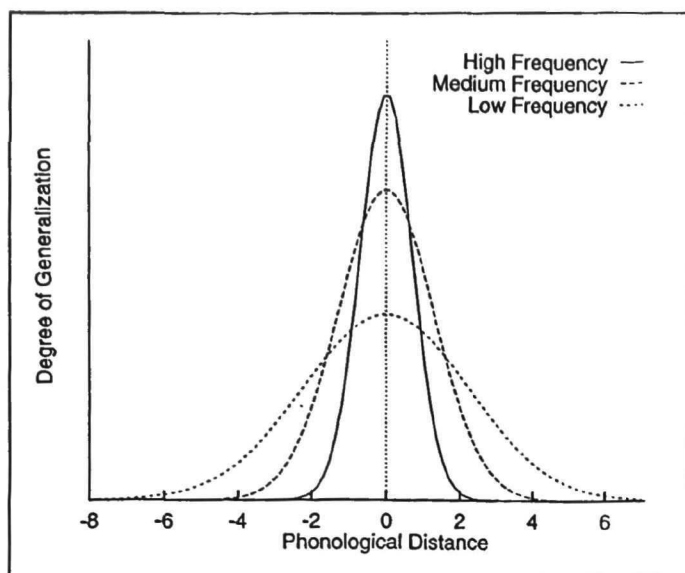
Figure 1: Effects of frequency and distance on generalization.

to change. Bybee relates this to morphological generalization through the idea of lexically based schemas, which emerge from lexical connections between lexical items. Sets of words with similar patterns of semantic and phonological connections reinforce each other to create emergent general patterns or schemas to which novel items can be assimilated. The greater lexical strength of high token-frequency items leads to greater lexical autonomy, which is reflected in weaker lexical connections to other items, yielding only an extremely localized and specific schema, with the net result of depressed generalization.

The specifics of our results are not compatible with suggestions that regular productivity is based on an entirely open, general /ed/ schema (Bybee, 1995) which applies virtually independently of lexical distribution But they are compatible with the idea that higher frequency leads to greater 'specificity' of lexical representation and thus a narrower neighborhood of potential generalization.

These issues clearly require further exploration, but again they suggest that future research might profitably move beyond the broad dual-route vs. single-route debate to more intricate models of lexically-based generalization.

Conclusions

We described two experimental studies which revealed similarity effects (experiments 1 and 2) and frequency effects (experiment 2) on regular generalization. These effects conflict with the dual-route account's claim that regulars are produced independently of the lexicon via a separate rule-route.

We have also found interactions between token frequency and phonological distance and proposed an intuitive account for this in the spirit of Bybee's (1995) proposals about lexical strength.

Finally, we have argued that Experiment 1 provides experimental evidence for product-oriented considerations in participants' treatment of nonce words ending in dental stops. These considerations conflict with the general stem+process orientation of the dual-route account, an orientation that it shares with current connectionist modelling of inflection. This seems, to us, to coincide with an increasing body of evidence for product-oriented considerations, both from theoretical analysis and from computational modelling, and to suggest a key direction for future work on inflection.

Acknowledgements

Reference

Bybee, J. (1995). Regular Morphology and the Lexicon. *Language and Cognitive Processes, 10*, 425–455.

Bybee, J. & Moder, C. (1983). Morphological Classes as Natural Categories. *Language, 59*, 251–170.

Clahsen, H., Eisenbeiss, S., & Sonnenstuhl-Henning, I. (1997). Morphological structure and the processing of inflected words. *Theoretical Linguistics, 23*, 201–250.

Daugherty, K. & Seidenberg, M. (1992). Rules or connections: The past tense revisited. In *Proceedings of the Fourteenth Annual Meeting of the Cognitive Science Society*, pp. 259–264. Hillsadle, NJ: Erlbaum.

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

Hahn, U. & Nakisa, R. German Inflection: Single- or Dual-Route?. submitted.

Hahn, U., Nakisa, R., & Plunkett, K. (1997). The Dual-Route Model of the English Past-Tense: Another Case Where Defaults Don't Help. In *Proceedings of the GALA '97 Conference on Language Acquisition*.

Jaeger, J., Lockwood, A., Kemmerer, D., Valin, R. V., Murphy, B., & Khalak, H. (1996). A Positron Emission Tomographic Study of Regular and Irregular Verb Morphology in English. *Language, 72*, 451–497.

Köpcke, K. (1993). *Schemata bei der Pluralbildung im Deutschen*. Tübingen: Gunter Narr.

Lee, B. (1996). On the processing of regular and irregular inflections: the symbolist-connectionist debate revisited. In Koster, C. & Wijnen, F. (Eds.), *Proceedings of the Groningen Assembly on Language Acquisition*.

MacWhinney, B. & Leinbach, J. (1991). Implementations are not conceptualizations: Revising the Verb Learning model. *Cognition, 40*, 121–157.

Marcus, G., Brinkmann, U., Clahsen, H., Wiese, R., Woest, A., & Pinker, S. (1995). German Inflection: The Exception that Proves the Rule. *Cognitive Psychology, 29*, 189–256.

Marcus, G., Pinker, S., Ullman, M., Rosen, T., & Xu, F. (1992). Overregularization in language acquisition. *Monographs of the Society for Research on Child Development, 57*.

Nakisa, R. C., Plunkett, K., & Hahn, U. (1998). A Cross-linguistic Comparison of Single and Dual-Route Models of Inflectional Morphology. In Broeder, P. & Murre, J. (Eds.), *Cognitive Models of Language Acquisition*. Cambridge, Mass: MIT Press *in press*.

Nakisa, R. & Hahn, U. (1996). Where Defaults Don't Help: the Case of the German Plural System. In *Proceedings of the 18th Annual Meeting of the Cognitive Science Society*. Mahwah, NJ: Erlbaum.

Penke, M., Weyerts, H., Gross, M., Zander, E., Münte, T., & Clahsen, H. (1997). How the brain processes complex words: An ERP study of German verb inflections. *Cognitive Brain Research, 6*, 37–52.

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

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

Plunkett, K. & Marchman, V. (1991). U-shaped learning and frequency effects in a multi-layered perceptron: Implications for child language acquisition. *Cognition, 38*, 43–102.

Plunkett, K. & Marchman, V. (1993). From rote learning to system building. *Cognition, 48*, 21–69.

Rumelhart, D. & McClelland, J. (1986). On learning past tenses of English verbs. In McClelland, J. & Rumelhart, D. (Eds.), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, Vol. Vol 2: Psychological and Biological Models. Cambridge, MASS: MIT press.

Seidenberg, M. & Hoeffner, J. (1998). Evaluating Behavioral and Neuroimaging Data on Past Tense Processing. *Language, 74*, 104–122.

Sereno, J. & Jongman, A. (1992). The processing of inflectional morphology in English. Paper presented at the Fifth Annual CUNY Conference on Sentence Processing, March 19-21.

Smolensky, P. (1988). On the proper treatment of connectionism. *Behavioral and Brain Sciences, 11*, 1–74.

Ullman, M. (1993). *The computation of inflectional morphology*. Ph.D. thesis, Dept. of Brain & Cognitive Sciences, Massachussets Institute of Technology.

Xu, F. & Pinker, S. (1995). Weird Past Tense Forms. *Journal of Child Language, 22*, 531–556.