Why will rat's go where rats will not?

Jenny Hayes[1], Victoria Murphy[1], Neil Davey[2], Pam Smith[1], Lorna Peters[2]

Departments of Psychology[1] and Computer Science[2], University of Hertfordshire, College Lane, Hatfield, AL10 9AB, United Kingdom.

Experimental evidence indicates that regular plurals are nearly always omitted from English compounds (e.g., rats-eater) while irregular plurals may be included within these structures (e.g., mice-chaser). This phenomenon is considered to be good evidence to support the dual mechanism model of morphological processing (Pinker & Prince, 1992). However, evidence from neural net modelling has shown that a single route associative memory based account might provide an equally, if not more, valid explanation of the compounding phenomenon.

1.Introduction

1.1 The Compounding phenomenon

Psycholinguistic research has shown that English compound words with irregular plural nouns in first position (e.g. mice-eater) are produced far more frequently than compound words with regular plural nouns in first position (e.g. *rats-eater) (Gordon, 1985). This paper outlines the standard explanation of this phenomenon based on Pinker and Princes's dual mechanism model (1992) but argues that an associative memory based explanation, which is explored using three connectionist models, might provide a more satisfactory explanation.

1.2 The Dual Mechanism Model's Explanation of Compounding

The dual mechanism model (Pinker & Prince, 1992), proposes that irregular nouns and their plurals are stored as memorised pairs of words in the mental lexicon (e.g. mouse-mice) but that regular plurals are produced by the addition of the [–s] morpheme to the regular stem at a post lexical stage (e.g. rat + s = rats). Compounds are created in the lexicon by joining two stems together to form one word. Thus as irregular plurals are stored in the lexicon they are available to be included within compound words. However, as only the singular stems of regular nouns are stored in the lexicon the plural form is never available to be included within compound words.

1.3 A Single Route Associative Memory Based Explanation of Compounding

An alternative explanation of this compounding phenomenon based on the frequency and patterns of occurrence of items in the linguistic input has not been explored fully. However an explanation of this sort may explain the treatment of both regular and irregular plurals in compounds (Murphy, 2000). Frequency counts of a sample of the CHILDES (Child Language Data Exchange System) corpora (McWhinney & Snow, 1985) have shown that the plural [-s] morpheme is a perfect predictor of word finality and furthermore, the plural [-s] morpheme is never followed by a second noun. Importantly, the reverse pattern is found with the possessive [-'s] morpheme since it is always followed by a second noun. Therefore, it might be that a noun rarely follows

the regular plural [-s] morpheme (i.e. patterns such as "*rat[s] chaser" do not occur) because the pattern "noun – morpheme [-s]- noun" is reserved for marking possession (such as rat's tail). Interestingly in other languages that do not have this competition between the plural and possessive morpheme such as Dutch (Schreuder, Neijt, van der Weide & Baayen, 1998) and French (Murphy, 2000), regular plurals are allowed within compounds. Irregular plurals may, however, appear in English compounds as they are not formed by the addition of the plural [-s] morpheme. Thus, irregulars do not compete with the possessive structure and as such may be followed by a second noun in a compound.

An associative memory-based account of inflectional morphology has been investigated in numerous connectionist models. Several models have successfully simulated the putative dissociation between regular and irregular inflection for both verbal morphology (Daugherty & Seidenberg, 1994) and plural morphology (Plunkett & Juola, 1999) using a single learning mechanism and no explicit rules. Furthermore, as well as being able to learn mappings from input to output, connectionist models have also been able to learn sequential mappings (Elman 1990). Thus it is predicted that a single route associative memory system could learn that the inclusion or omission of the regular plural morpheme [-s] is influenced by where that [-s] morpheme occurs in a sequence of language input. Three neural net models are considered here. The first investigates any role that [-s] (whether plural or possessive) might play as a predictor of word finality. The second and third models analyse whether learning about the word that follows an [-s] morpheme is sufficient to drive learning about compound formation in English.

2. Neural net modeling

2.1 Experiment 1.

Experiment 1, was designed to test the degree to which [s] indicates word finality in a stream of concatenated letters. A neural network was trained on a concatenated stream of 200 sentences of child directed speech taken from CHILDES (MacWhinney & Snow, 1985). A word-ending marker was attached to each word and the words (including a word-ending marker) were concatenated to form a stream of 3596 letters. Each letter was encoded using one of 26 random 5-bit vectors (one for each letter in the alphabet). The word-ending marker was encoded using a 27th 5-bit vector. The network was required to predict the next letter it expected to occur given the letters it had seen previously. The network consisted of 5 input units, 30 hidden units, 5 output units and 35 context units. The network was fully recurrent so that at any point in time the state of the hidden units and the output units at the previous time step were used as additional input (Elman, 1990). It was hypothesised that on a next letter prediction task of this kind, a neural network would learn that after the input [-s] there was a high probability that the next input would be a word ending marker

Test Sets and Results: As predicted, at the beginning of a word the error was high but as more letters were presented to the network the error decreased until it was at its lowest at the end of the word. The network's ability to learn that [–s] is a good predictor of word finality was tested using 19 unseen words that ended in [-s] and 19 unseen words that ended in other letters. The network was more accurate (i.e. the error was lower) at predicting a word ending marker after an [-s] than after all other

letters combined. This simulation was completed to confirm that a model with a single learning mechanism and no explicit rules, trained on child directed speech, could learn that after [-s] there was a high expectancy that the next item would be a word-ending marker. This overwhelming pattern of [-s] at the end of a word may influence language learners to omit [-s] from the middle of words such as compounds.

2. 2. Experiment 2.

The aim of this experiment was to examine how highly consistent patterns in the input (i.e. that a plural noun is never followed by another noun while a possessive noun is always followed by a second noun) might drive learning about how to manipulate plurals within noun-noun compounds. The network was required to predict the next word it expected to occur given the words it had seen previously. It was impossible for the network to predict the exact word that followed in the input. However, the network was expected to learn which syntactic category the next item would come from. Thus the network was expected to make a first order distinction between the function of nouns and verbs, determiners and adjectives (Elman, 1990). Furthermore from these induced syntactic categories the network was expected to learn a second order distinction that only "verbs" could appear after some [-s] morphemes and only "nouns" could appear after other [-s] morphemes. It was impossible for the network to distinguish between the possessive and the plural [-s] as both were encoded in exactly the same manner in the input. However, the network was trained on one group of words that were represented as having the properties of possessives, plurals and singulars, a second set was only represented as singulars and plurals and a third group was only represented as singulars and possessive. It was predicted that the tokens making up these three groups of words would cluster together in the hidden layer representations. The network was trained on a concatenated stream of 2000 legitimate English sentences constructed from a lexicon of 38 words. A sentence-ending marker was attached to each sentence and the sentences (including the sentence-ending marker) were concatenated to form a stream of 14,600 words. Each word (including the sentence-ending marker) was encoded using a 39-bit localist coding scheme. The presence or absence of [-s] at the end of a word was also explicitly coded. A simple recurrent network was used so that at any point in time the state of the hidden units at the previous time step were used as additional input (Elman, 1990).

Results: Figure 1, shows a typical representation of the first two principle components of the hidden unit representations. The dotted line superimposed on the PCA diagram shows the divide between the way nouns and verbs are represented in the hidden units. It is also apparent that the network has also represented determiners and adjectives separately. Most interestingly, nouns which were included in the training set as both "plurals and possessives", items that were only included as "possessives" and items which were only included in the "plural" form are all represented separately within the cluster of words ending in [-s]. Therefore, Experiment 2 showed that a neural net was able differentiate the plural and possessive [-s] depending on the words which followed it in the input even though the two types of [-s] had exactly the same encoding characteristics.

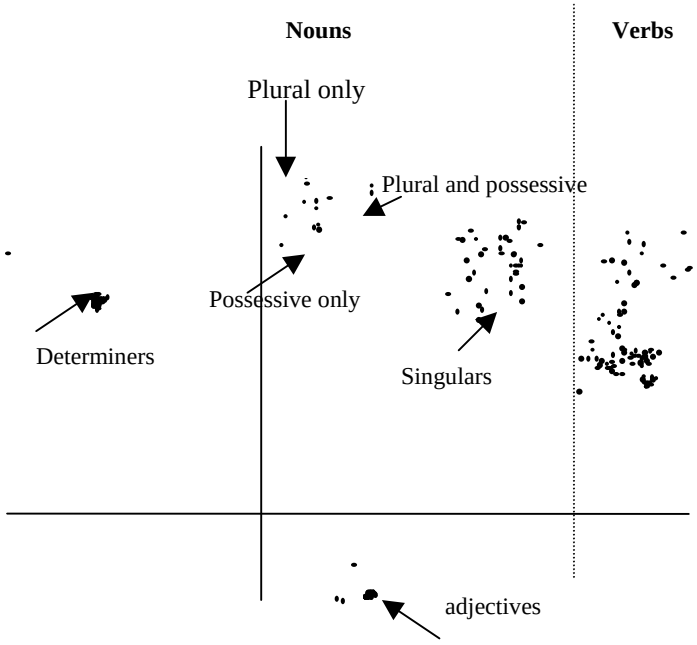

Figure 1. First two principle components of the hidden layer representations in Experiment 2

2. 3. Experiment 3.

In Experiment 2, the network was able to group nouns that in the training set were behaving as "plural and possessive" or as "plural" or "possessive" only. However, the network could not totally disambiguate plurals from possessives. In this third simulation, the network that was used in Experiment 2 was amended to include an extra input unit that encoded whether the subject of the sentence in which the word occurred was either a plural or a singular noun. Hence, although both "plural" and "possessive" words were coded as ending in [-s], only plural items were encoded as ending in [-s] and being plural as possessive words were encoded as ending in [-s] but being singular. The same training set and task utilised in Experiment 2 was employed. It was predicted that with the addition of this minimal semantic information the network would be able to disambiguate "plural" nouns from "possessive" nouns. It was predicted that in the hidden units the plural and possessive nouns would be represented separately.

Results: Figure 2, shows a typical representation of the first two principle components of the hidden unit representations. From the PCA it is evident that once again nouns and verbs determiners and adjectives are represented separately in the

hidden units. With the addition of the semantic information it is now evident that singular, plural and possessive nouns are all represented separately. Interestingly, both plurals and singulars i.e., items that may be followed by a verb lie in similar positions on the x axis, while the possessives are clustering with adjectives i.e., with other items that are followed by nouns. Therefore, Experiment 3 shows that learning about the different functions of the [-s] morpheme is enhanced with the addition of the very minimum of semantic information.

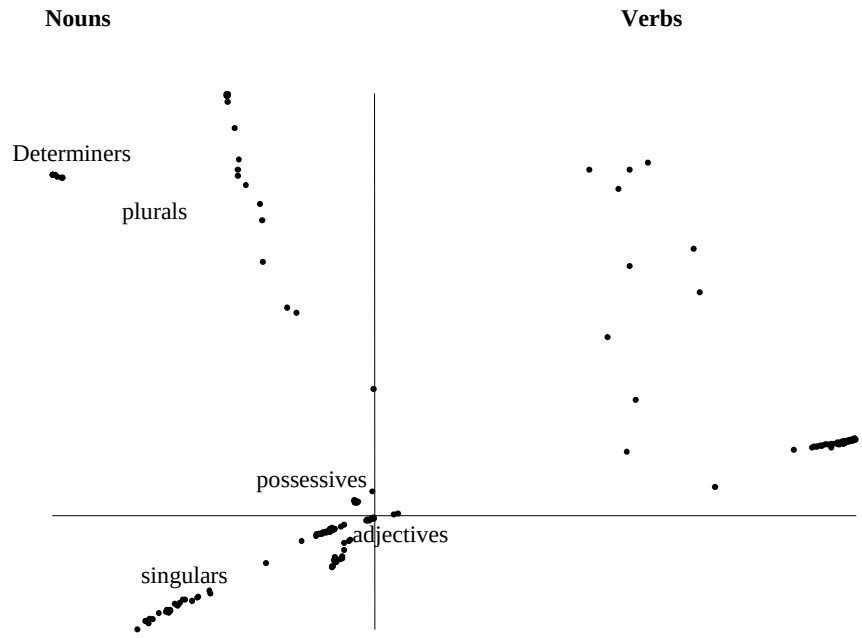

Figure 2. First two principle components of the hidden layer representations in Experiment 3

3. Discussion

From Experiment 1, it is evident that a neural net model trained on child directed speech was able to learn that [-s] is associated with word-finality. This overwhelming pattern of [-s] at the end of words might influence language learners to omit [-s] from the middle of words such as compounds. Experiment 2, showed that the net was able to learn that [-s] followed by one set of words was different from [-s] followed by a different set of words even though the [-s] was encoded in exactly the same way in the input. The same might be true for the language learner. Both the possessive [-s] and the plural [-s] sound the same phonetically but the patterns in which the two different types of morpheme appear in the input may be sufficiently distinct as to indicate that one type of morpheme performs a specific linguistic function and the other performs another type of linguistic function. From Experiment 3, it is evident that learning that the plural and possessive morphemes are only legal in certain sequences may be refined as the child learns that semantically, the plural morpheme refers to many things while the possessive morpheme usually refers to one thing. These three models taken together provide evidence for an associative account of

compounding. In this associative account, the language learner is sensitive to the fact that the [-s] morpheme tends to nearly always occur at the end rather than in the middle of a word (Experiment 1). Furthermore, simply by exposure to the [-s] morpheme (i.e. without the plural or the possessive [-s] morpheme being explicitly labelled as being different from each other), the language learner is sensitive to the fact that the same [-s] morpheme occurs in different patterns in the input (Experiment 2). With the addition of the absolute minimum of semantics, namely the numerical context in which the phrase is uttered, the language learner seems able to differentiate between the plural and the possessive morpheme (Experiment 3). The possessive morpheme may be followed by a second noun but the plural morpheme may not be followed by a second noun. When faced with a noun-noun compound the language user may delete the plural morpheme from the end of the first noun not because regular items of morphology are different in kind from irregulars and represented as "rules" in the brain but simply because this pattern is used to denote possession not plurality. Thus the dissociation between the treatment of regular and irregular morphology in compounds may result from the fact that one type of morphology is subject to competition with the possessive morpheme but the other is not. As this alternative hypothesis is explored further, it may become apparent that this plural dissociation in compounds is not good evidence to support the dual-mechanism model.

References

1. Daugherty, K. G. & 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.
2. Elman, J. L. (1990). Finding structure in time. *Cognitive Science, 14*, 179-211.
3. Gordon, P. (1985). Level-ordering in lexical development. *Cognition, 21* , 73-93.
4. MacWhinney, B. & Snow, C. E. (1985). The Child Language Data Exchange System. *Journal of Child Language, 12,* 271-296.
5. Murphy, V. A. (2000). Compounding and the representation of inflectional morphology. *Language Learning, 50,* 153-197.
6. Pinker, S. & Prince, A. (1992). Regular and irregular morphology and the psychological status of rules of grammar. In L. A. Sutton, C. Johnson, & R. Shields (Eds.), *Proceedings of the 17th Berkeley Linguistics Society* (pp.230-251).
7. Plunkett, K. & Juola, P. (1999). A connectionist model of English past tense and plural morphology. *Cognitive Science, 23,* 463-490.
8. Screuder, R., Neijt, A., van der Weide, F. and Baayen, R.H., (1998). Regular plurals in Dutch compounds: Linking graphemes or morphemes? *Language and Cognitive Processes*, 13, 551-573.