The /s/ morpheme and the compounding phenomenon in English

Jenny Hayes[1] (J.Hayes@herts.ac.uk)
Victoria Murphy[1] (V.A. Murphy@herts.ac.uk)
Neil Davey[2] (N.Davey@herts.ac.uk)
Pamela Smith[1] (P.M.Smith@herts.ac.uk)
Lorna Peters[2] (L.Peters@herts.ac.uk)
Departments of Psychology [1] and Computer Science [2], University of Hertfordshire, College Lane, Hatfield, AL10 9AB, United Kingdom.

Abstract

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). Using empirical evidence and neural net modelling, the studies presented here demonstrate how a single route, associative memory based account might provide an equally, if not more, valid explanation of this phenomenon than the standard dual mechanism based theory (Marcus, Brinkmann, Clahsen, Weise & Pinker, 1995).

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).

1.2 The Dual Mechanism Model's Explanation of Compounding

The dual mechanism model (Pinker, 1991), 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. 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 (Marcus et al, 1995).

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. This polyfunctionality of the /s/ hypothesis is explored here using three neural net simulations and an empirical study.

2. Neural Net Modeling

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 whether the presence of /s/ predicts 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.

2.1 Experiment 1.

Experiment1, 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. The network was trained on 200 passes through the sequence of 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. A fully recurrent and a SRN (Elman, 1990) architecture were tested and produced similar results. 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 Set 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 ($t=-2.08$, $df=18$, $p<0.05$). 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. Interestingly, /s/ is associated with word finality even though /s/ appears in the middle of numerous words. 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 possessives. It was predicted that the tokens making up these three groups of words would cluster together as three distinct sets 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. The network was trained on 10 complete passes through the sequence of 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 principal 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. 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 to make some differentiation between the plural and possessive /s/ depending on the words that followed it in the input even though the two types of /s/ had exactly the same encoding characteristics.

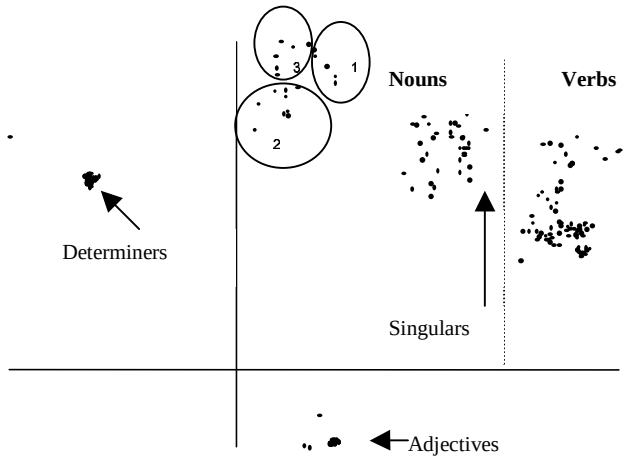

Figure 1. First two principal components of the hidden layer representations in Experiment 2. Area 1 corresponds to the representational area occupied by items that appeared in the context of both plurals and possessives. Area 2 corresponds to the representational area occupied by items that appeared only in a possessive context and area 3 corresponds to the representational area occupied by items that appeared only in the plural context.

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 from Experiment 2 was employed. With the addition of this minimal semantic information, the network is expected 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 principal components of the hidden unit representations. From the PCA it is evident that once again nouns, 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. Singular and possessive nouns (both of which are actually singular nouns) are located in a similar position but plural nouns are now represented quite 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

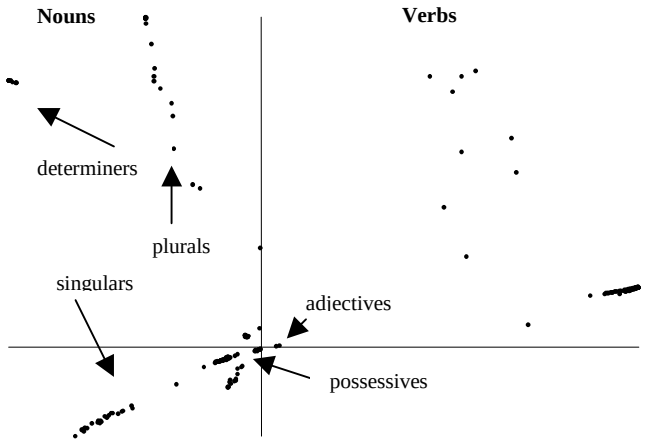

Figure 2. First two principal components of the hidden layer representations in Experiment 3

2.4 Discussion of Neural Net Modelling

From Experiment 1, it is evident that a neural net model trained on child directed speech was able to learn that /s/ is strongly associated with word-finality (even though /s/occurs frequently in the middle of words). This overwhelming pattern of /s/ at the end of words might influence English 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.

Table 1: Examples of Compounds used as stimuli in Experiment 4.

Group	Example of contextualising sentence	Examples of compounds
(1) Possessive nouns	Last week, I left my purse in a London taxi. Luckily, I managed to signal to the	taxi's driver
(2) Regular plural nouns	I feed four cats, a Burmese, a Siamese and two lovely old Persians. I enjoy being a	cats feeder
(3) Irregular plural nouns	Women always get lowly jobs. In the nursery rhyme the farmer's wife is nothing more than a	mice chaser
(4) Comparatives or superlatives	Greg is very modest. He was amazed to hear that his song is still the record company's	biggest seller
(5) singular nouns ending in phoneme /s/	We'll have a larger lawn and mowing the grass will take longer. I'm thinking of employing a	grass cutter
(6) singular nouns ending in a phoneme other than /s/	Stephen is so skilled at mixing cocktails that the hotel want him to work permanently as a	drink server

3. Experiment 4:Compound Processing Study

The compounding phenomenon was further tested by asking 22 native adult English speakers to process "noun-noun" compounds as part of an "on-line" lexical decision (LD) task. This is important as most research has focussed on production (e.g., Gordon, 1985; Murphy, 2000). In this experiment, participants were required to categorise 216 compounds as being words or non-words having seen them presented visually on a computer screen and proceeded by a contextualising sentence. The mean response time for processing different types of compound (see stimuli) was examined in a within subjects design. Two hypotheses were tested to examine the associative explanation of compounding. These were:- (1) If, as the first neural net simulation (Experiment 1) confirmed, language users associate /s/ with word finality, then compounds in which the first noun ends in /s/ should be processed more slowly than compounds that do not include a first noun ending in /s/. (2) If, as the polyfunctionality hypothesis indicates, /s/ appearing in the middle of a compound made up of two nouns is interpreted as indicating possession rather than plurality, then compounds containing possessive nouns should be processed more quickly than compounds containing plural nouns**.**

Two further hypotheses were investigated to test the dual mechanism model's explanation of compounding. (1) Pinker (1991) stated that:

> "because it categorically distinguishes regular from irregular forms, the rule-association hybrid predicts that the two processes should be dissociated from virtually every point of view…..[including] reaction time ……. " (p 253).

However, the dual mechanism model makes no directional prediction as to which type of morphology might be processed more quickly. Beck (1997) asked native adult English speakers to supply the past tense of a series of present tense regular and irregular verbs. Beck found that both low (mean response time 477 msec) and high (mean response time 508 msec) frequency regulars were produced more quickly that both low (mean response time 581) and high (mean response time 535 msec) frequency irregulars. By collecting reaction times in Experiment 4, it was possible to test the speed at which the two types of morphology were processed within compounds in a lexical decision task. The following two hypotheses were tested, (1) compounds containing irregular morphology and compounds containing regular morphology will be processed at different speeds (2) more specifically, compounds containing irregular plurals and compounds containing regular plurals will be processed at different speeds.

3.1 Stimuli

The first noun in each compound was taken from one of six groups. These were: - (1) possessive nouns (2) regular plural nouns (3) irregular plural nouns (4) comparative or superlatives (5) singular nouns which ended in phonetic /s/ (6) singular nouns which ended in a phoneme other than /s/. Each group of first nouns were matched for frequency. The second noun in each compound was a deverbal noun, i.e., a noun that is formed from a verb (e.g.s, walker, chaser). All second nouns were matched for frequency. Table 1.shows examples of each type of compound tested. The apostrophe was omitted from all the possessive nouns thus making it impossible to

distinguish between the plural and possessive nouns (cf. the neural net used in Experiment 2). However, each compound was preceded by a contextualising sentence, (cf. the neural net used in Experiment 3) which pilot work had shown would lead the first noun in the compound (e.g., rats in *rats eater) to be interpreted appropriately. To ensure uniform treatment of all stimuli, contextualising sentences also preceded the first noun even where they were not taken from the plural or possessive groups (see Table 1. for examples of contextualising sentences).

3.2 Results

Table 2. Mean reaction times

	Mean reaction time in milliseconds	Standard deviation of Mean reaction time	Difference between means in milliseconds
Comparisons to test the associative account			
All groups ending in /s/	1285	465	
Final phoneme other than /s/	1205	455	80*
Regular plurals	1277	492	
Possessives	1191	437	86*
All groups containing regular morphology	1231	450	
Irregular plurals	1339	470	108*
Regular plurals	1277	492	
Irregular plurals	1339	470	62
Comparison to test the time difference between processing plurals and processing other types of morphology			
Regular and irregular [illegible]	1291	479	
All other items of morphology	1188	424	103*

* Difference reliable at alpha = 0.05

Mean reaction times in milliseconds are shown in Table 2. Two planned comparisons were conducted to test the associative explanation of compounding. Firstly, compounds with a first noun ending in /s/ were processed more slowly than compounds where the first noun ends in another letter (a mean difference of 80 milliseconds) ($t=4.41$, $df=21$, $p<0.05$). It took participants an average of 86 milliseconds longer to process compounds containing regular plurals than compounds containing possessive nouns ($t=2.195$, $df=21$, $p<0.05$). These two findings support the outcomes of the neural net simulations in Experiments 1,2 and 3. Two planned comparisons were conducted to test the dual mechanism model's explanation of compounding. All types of regular morphology were processed more quickly that irregular plurals (mean difference of 108 milliseconds) ($t=3.22$, $df=21$, $p<0.05$). It took participants an average of 62 milliseconds longer to process irregular plurals than regular plurals but this difference was not found to be reliable. A post hoc analysis was also conducted to test the difference between the time taken to process compounds containing both types of plural and the time taken to process compounds containing other types of morphology (mean difference of 103 milliseconds). A Tukey's HSD test found this difference to be reliable ($F=11.29$, $df=21$, $p<0.05$).

3.3 Discussion

Consider first the two hypotheses that tested the associative explanation of compounding. It took longer to process compounds in which the first noun ended in /s/ than compounds which did not include a first noun ending in/s/. Language users, like the network in Experiment 1, found it harder to process /s/ in the middle of a word. Furthermore, possessive nouns are easier to process than plural nouns in the middle of compounds even though they share exactly the same phoneme. The /s/ morpheme in the middle of a word seems to indicate possession not plurality. Consider next the two hypotheses that tested the dual mechanism's explanation of compounding. Similar to Beck's (1997) production data, it took participants in this experiment less time to process all types of regular morphology than it took them to process irregular plurals (the only type of irregular morphology tested). However, no difference was found in the time taken to process regular and irregular plurals, despite Pinker's (1991) prediction that the two types of morphology should be dissociated "from virtually every point of view" (p 253). Interestingly, a reliable difference was found when reaction times to both types of plural were collapsed together and compared with reaction times to comparatives and superlatives and possessives (all items of regular morphology) collapsed together. Adult language users find it relatively difficult to process either type of plural in the middle of compounds. However, contrary to the predictions of the dual mechanism model, adults seem to have no difficulty processing other items of regular morphology (i.e., items which are produced at a post-lexical stage) within compounds (cf. Marcus et al, 1995). It has been argued elsewhere, that due to the competition with the possessive structure, language users omit regular plurals from compounds. Furthermore, guided by this template, i.e., that plurals do not occur within compounds, mature language users also begin to

omit irregular plurals from compounds, (Hayes, Smith & Murphy, unpublished manuscript)

4. General discussion

Experiment 1 showed that /s/ is associated with word finality. Furthermore, participants in the empirical study took longer to process compounds which contained /s/ than compounds that did not contain /s/ (regardless of what type of /s/ it was). Both strands of evidence would seem to indicate that /s/ is linked to word finality despite the fact that /s/ occurs frequently in the middle of many words. 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. Evidence from Experiment 2, showed that a neural network 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. From Experiment 3, it was 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. The empirical evidence also showed that one type of /s/ morpheme was processed more quickly within compounds than another type of /s/ morpheme, although it was denoted by same phoneme. 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 seem to be sufficiently distinct to indicate that one type of morpheme performs a specific linguistic function and the other performs another type of linguistic function. As well as providing support for the associative account Experiment 4 also calls into question the dual mechanism model's explanation of compounding. No difference was found between the time taken to process regular and irregular plurals. However, participants were able to process some items of regular morphology within compounds relatively quickly. It seems to be plurals (of either kind), rather than items of regular morphology, that adults find difficult to process within compounds.

The three models taken together with the empirical work provide evidence for an associative account of compounding. In this associative account the language user relies on properties of the linguistic input itself and not on distinct ways of representing "rules" versus associations to drive linguistic behaviour. More specifically in the case of compounding, the language user learns that the /s/ morpheme tends to nearly always occur at the end rather than in the middle of a word. Furthermore, the language learner is sensitive to the fact that the same /s/ morpheme occurs in different patterns in the input. 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. 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 not because regular items of morphology are different in kind from irregulars and represented as rules 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 form 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

Beck, M-L., (1997). Regular verbs, past tense and frequency: tracking down a potential source of NS/NNS competence differences. *Second language Research, 13,* 93-115.

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.

Elman, J. L. (1990). Finding structure in time. *Cognitive Science, 14,* 179-211.

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

Hayes, J.A., Smith, P.M., Murphy, V.A. (Unpublished manuscript). Modality effects in compounding with English inflectional morphology.

Marcus, G.F., Brinkmann, U., Clahsen, H., Weise, R. & Pinker, S., (1995). German inflection: The exception that proves the rule. *Cognitive Psychology, 29,* 189-256.

MacWhinney, B. & Snow, C. E. (1985). The Child Language Data Exchange System. *Journal of Child Language, 12,* 271-296.

Murphy, V. A. (2000). Compounding and the representation of inflectional morphology. *Language Learning, 50,* 153-197.

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

Plunkett, K. & Juola, P. (1999). A connectionist model of English past tense and plural morphology. *Cognitive Science, 23,* 463-490.

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.