Cognition, 7 (1979) 35–47

Brief Report

Remembering plurals: Unit of coding and form of coding during serial recall*

HUGO VAN DER MOLEN** and
JOHN MORTON

*MRC Applied Psychology Unit, Cambridge, England****

Abstract

Subjects were required to recall lists of six words which had been presented visually in sequence. One or two of the words might be plural nouns. A substantial number of errors occurred in which the plural form became detached from its original root. This is taken as evidence for a morpheme-based code as opposed to a unitary word code. A significantly high proportion of these errors maintained the form of the plural (/S/, /Z/ or (ə Z/) which is considered evidence for a primarily phonological coding of the plural morpheme. There is however a suggestion that on some occasions the plural is coded morphemically—i.e., in a way which does not distinguish between the various plural endings.

It is currently common to talk in terms of the code in which material is registered mentally in particular tasks. By means of a simple illustrative experiment we wish to make a distinction between two factors; the *units* on which the code is "based" and the *form* of the code. The distinction between the two can best be described by means of an example. The word CATS can be represented in a number of ways. There could be a single unit, which would be at the word size. Alternatively the code could be morpheme-based in which case there would be two units CAT + plural morpheme. A third possibility is that the characters are the only units. If the characters were letters then at this size of coding there would be no difference between the forms of representation of CATS, ACTS and SCAT.

Given that we have units of a particular size we then have options as to the *form* of the code. The plural morpheme could thus be represented mor-

*The research was carried out in 1974 while the first author was supported by a Graduate Fellowship from the Rotary Foundation.
**Now at the University of Groningen, The Netherlands.
***Requests for reprints should be sent to John Morton, M.R.C. Applied Psychology Unit, 15 Chaucer Road, Cambridge, CB2 2EF, England.

phemically, in which case the plural morphemes for CATS, DOGS and MEN will be identical. Alternatively the plural morpheme could be coded phonologically, in which case that for CATS would be represented as /s/ and that for DOGS as /z/. With a visual code the two have an identical form, namely 'S'.

In discussing the internal representation of words we will prefer to talk about the dominant unit of coding rather than trying to establish a particular unit as the only one. Equally we will see that in practice the form of the code may not be restricted to one type in the course of a single task. The task we will examine is that of serial recall where words were presented visually to our subjects. The units and form of the coding will be reflected in the type of errors found in recall. Our interest was focussed specifically on the way in which plural nouns are represented internally. There are a number of possible ways in which the plural might be coded and linked to the root morpheme.

Word-based and character-based coding

In the extreme case of word-based coding one would expect relatively few errors occurring in which the plural inflection becomes detached from its original root, or in which the separated plural inflection becomes attached to another root.

In the extreme case of character based coding the plural inflection is related to the rest of the word in just the same way as any other component. Thus, in a word like SPOTS, the final S would be linked to the preceding T in just the same way as the initial S was linked to the P. This could apply equally whether the code has a visual form or a phonological form. Thus errors of the SPOTS–SPOT type should be no more common than errors of the SPOTS–POTS type (given suitable controls for frequency etc). And in general one would expect omission errors to involve both parts of the word equally.

Morpheme-based coding

In this case the root and plural parts of the word are linked together in a way which is different from the intercharacter links in the root. Specifically, the former would be weaker. Thus we would expect recall errors of the form SPOTS–SPOT considerably more often than errors like SPOTS–POTS. We have a further decision to make concerning whether the plural part of the stimulus word can survive in memory separated from its root. If it can then we would expect there to be errors which look like the transfer of the plural from one stimulus to another. Thus: SPOTS CAT–SPOT CATS.

One study which examines the possibility of the morpheme as unit of coding is by Murrell and Morton (1974) who showed that the prevailing units involved in facilitation of visual recognition of words were the morphemes, rather than the words or characters. Taft and Forster (1975) came to similar conclusions with a lexical decision task. Within the framework of the logogen model (Morton, 1968, 1969, 1970) this would lead to the prediction of morpheme-based coding in the short term serial recall paradigm. Concerning the *form* of the coding there would be two main alternatives, given that the code was morpheme-based.

(a) Phonological code. In such a case the plural morpheme would be represented as one of the three alternative regular forms of the plural /s/, /z/ and /ðz/ or the feature list which correspond to these. Three equivalent classes of noun are formed as a function of the final phoneme in the root noun. Thus /ðz/ appears after sibilants and affricates as in 'horses, judges', /z/ appears after all other voiced phonemes as in 'dogs, windows', and /s/ after all other voiceless phonemes as in 'cats, mops'. The same symbol would be used to represent the endings of 3rd person singular present tense verb forms such as "teaches, opens, sits".

(b) Morphemic code. Here all plurals would be coded in exactly the same way as, say, [PL]. So CATS would be coded as [CAT] + [PL], DOGS would be [DOG] + [PL], and MEN would be [MAN] + [PL]. The verb forms, however, would be coded differently, SITS being represented, perhaps, as [SIT] + [SING] + [3rd] + [PRES]. Note that the [CAT] component of CATS would be identical to that of CAT.

In principle we can decide between the alternatives (a) and (b) in a fairly simple way. We present the subjects with a list of words which contains a plural, say, BEDS, and three singular nouns, say CAT, DOG and HORSE. We are particularly interested in the occasions where BED is made as a response. In this case the plural has become detached from the root noun and we would suppose that it would be floating, as it were, in memory and could be available for recall with some other noun. The important question is whether or not the three classes of noun are equally likely to appear in plural form. If the plural is coded morphemically then CAT, DOG and HORSE are equally likely to appear in the plural form. If the code were phonological, however, the particular form of the plural, in this case /z/, would require that it went only to a noun of the correct class–in this case DOG. Thus by examining the fate of the lost plurals we can decide between phonological and morphemic forms of coding. Unfortunately, the predictions are not quite clear as it seems. If one allows the possibility of decay or interference during the recall inter-

val then a /z/ form could be changed into a /s/ form etc., in store, giving rise to an error which would appear to support the notion of morphemic coding.

Two further predictions can be made. If the plural is coded morphemically then the loss of a plural from a stimulus noun should never lead to an error on a verb such as producing KNOWS instead of KNOW. The suffix on a verb does not indicate a plural. If the coding were phonological, however, there would be no obstruction to such a transfer (on the assumption that the derivational history and semantic or syntactic force of a suffix is not maintained during those processes involving a phonological code). With a morphemic code there would be no obstruction to the formation of plurals from irregular forms (such as MEN for MAN). With a phonological code this would be impossible.

There is a little prior work which would lead to particular predictions other than the quite generally reported finding that phonological confusions are found in material presented visually. As it currently stands the logogen model is indifferent as to morphemic or phonologically coded plurals in this paradigm.

Method

Stimulus Materials

The stimulus lists each consisted of six words. They can best be seen as being made up of a set of three test words and a set of three 'context' words. The words in the test sets were all monosyllabic singular nouns. Each set contained one noun from each of the main plural-forming classes. The three words in each set were chosen to have roughly the same frequency of occurrence as in Kucera and Francis (1967). An additional constraint was that the plurals of the three nouns in a test set should also have the same frequency of occurrence. The sets of *context* words contained either 2 (in 6 cases), 1 (in 3 cases) or no plurals (in 3 cases). Such plurals were always regular and where there were two plurals they were of the same phoneme type. The three types of plural were equally represented, each being found in three sets. The rest of the context sets were made up of singular nouns which take irregular plurals, 3rd person plural present tense verbs, adjectives and prepositions. The sets are given in the Appendix.

In the stimulus lists one test set and one context set were combined so that the three test words followed each other in the beginning, the middle or the end of the list. The place of the plurals in the lists varied but was balanced across the plural types. Different combinations of test and context sets were used for the six subject groups tested. Subjects, all adult female volunteers, were tested in groups of about 15 each. A total of 86 were tested.

Stimulus Presentation

The words were projected from slides—white letters on a black field, and were presented at a rate of 1.7 sec (1 sec exposure and 0.7 sec between slides). Following the last word in a list a dim, blank slide was presented as a sign for the subjects to write down their responses. About 30 sec were allowed for the responses. Each list was written on a separate page in a booklet. Two practice lists were given, after each of which the correct responses were written on a blackboard to familiarise the subjects with the procedure. Neither of these lists contained plurals and no mention of plurals was made during the instruction period. There followed the twelve test lists which were presented once through with, again, 30 sec for written responses. After the experiment, subjects were asked to write down their comments on the experiment, notes on any strategies they had used or difficulties they had encountered. Only one of the subjects made any mention of the plurals, to the effect that they were more difficult. Other comments which were made include that the similarity of succeeding words made recall more difficult, that alliteration helped; that function words were more difficult than nouns. We were satisfied that the purpose of the experiment was not apparent to the subjects and that we could be confident that no special strategies were being used.

Results and Discussion

Recall errors were only scored with respect to the plurals. For scoring purposes we refer to the root upon which the plural was formed as the *plural word* or *W*, and to the plural morpheme itself as *P*. Either or both of these components could be recalled. When the word, *W*, is omitted *P* can still be recalled attached to another word which may be either one which was in the list (either one of the test set or the context set) or an intrusion. In those lists containing two plurals it would be possible for either or both to be recalled with or without the plural. In the cases where *P* is attached to some word other than *W*, that word may belong to the same or a different phonological class as the stimulus *W*.

Word-, Morpheme- or Character-based coding

We can first summarise the evidence that the coding is strongly morphemic. There are two kinds of event which indicated that the root and the plural morpheme have become separated.

1. Loss of the plural: Of the 1290 plural nouns presented to the subjects, 406 were unambiguously forgotten (i.e., not counting possible replacements

such as *faces* for phrases). Of the 884 nouns recalled 150 were recalled in the singular form. We take this 17% of errors as an indication of the degree of independence of the plural morpheme. The alternative explanation of the omission is that it was the final character (letter of phoneme) which was omitted rather than the plural morpheme. This would imply that any other final character would have the same status as the plural and be equally likely to be omitted. By simple extension the initial character might be supposed to be equally vulnerable.

(a) The probability of an /s/ being omitted from the beginning of a stimulus word. There were only two possible cases—*spaces* which was reproduced as *paces* by only one of the 86 subjects, and *sit* which was never reproduced as *it*.

(b) The probability of other final phonemes being omitted. A number of possibilities occurred in the stimulus lists. Any of *guest, court, lawn, warm, tooth, wall, write, team* and *lamp*, could lose the final phoneme and still produce an English word. In the data only two examples were found with TEA substituted for TEAM and LAMB for LAMP.

(c) The probability of other initial phonemes being deleted. From the set of possibilities *train, price, bridge, queen* and *fruit* only FRUIT → ROOT was found from one subject.

None of these baselines were designed for. However, the number of occasions on which evidence leads us to suppose the coding to be character based is very low indeed, less than 1% of the possible occasions.

2. On a total of 150 of the 1032 lists plurals were added to one or two of the test set nouns. The probability of such an addition was heavily dependent upon whether or not a plural had been in the stimulus list and whether or not an error had been made in the recall of the plural nouns. The figures are given in Table 1 (a, b). Firstly it is clear that the likelihood of a plural being added is very small in the case that no plural had been in the stimulus list. That is we are not seeing a generalized tendency to add plurals to the list regardless of the stimulus. Secondly, the probability of there being a spare plural is much greater if an error had occurred on a plural noun. It didn't make more difference whether the error had been the omission of a plural from a remembered stem or the omission of the stem itself. In both cases, then, we can imagine that unattached plural morphemes were available in the appropriate storage system.

A further check that these additions were plural rather than characters is given by the probability of an /s/ being added at the beginning of a word. It would, of course, be necessary for the result also to be an English word. Such a control was not designed for but the stimulus words included *lake, park, team, tops* and *inch* which could be turned into *slake, spark, steam, stops*

Table 1a.

		No. of lists	No. of lists with plurals added in response	%
All stimulus lists with plurals		774	141	18.2
No plurals in stimulus list		258	9	3.5
	Total	1032	150	

Table 1b. *(for stimulus lists with plurals)*

	No. of lists	No. of lists with plurals added in response	%
Response error in a plural noun	507	113	22.3
No response error in a plural noun	267	28	10.5

and *cinch* (the latter with a phonological code only, the others with either phonological or visual). None of these errors occurred. Using this data as a control does involve some minimal extension to the underlying model but not excessively so.

Therefore we confidently conclude that the unit of coding of the material during the recall interval is indeed strongly morpheme-based, in that the linkage between the root and the plural is weaker than the links within the root.

A recent paper by Glanzer and Razel (1974) showed that free recall of a list of 15 bimorphemic nouns such as *chestnut, teaspoon, earthquake* was as good as recall of disyllabic single morpheme nouns like *donkey, satin, fibre* and slightly better than recall of monosyllabic words such as *elm, yawn, scrap.* Their interpretation is that the unit of storage is a word rather than a syllable or a morpheme.

The same conclusion was drawn by Osgood and Hussain (1972) who showed that no transfer occurred from compounds like STOCK MARKET to the subsequent recognition of MARKET.

The difference between this pair of results and ours lie in the nature of the complex words. We have used inflections, as did Murrell and Morton, whereas

the experiments of Glazer and Razel and Osgood and Hussain used nouns compounded from two free morphemes. It seems likely that the psychology needs two terms for the linguists' one, "morpheme" being adequate for the latter but not for our purposes.

Phonological vs Morphemic code of the Plural

Having decided that the plural is coded separately from the root morpheme we needed to decide the nature of the code. We pointed out in the introduction that if the plural is coded phonologically then having become detached from its root noun it could only reattach to a noun of the same type. If the plural is coded morphemically then it could be reattached to any other noun. The most extreme 'morphemic' code would be the subject just remembering something like "there was a plural noun somewhere, so let's just make one of the responses plural".

In the analysis we will deal separately with displaced plurals and duplicated plurals since they could in principle give different answers.

In analysing the fate of displaced plurals we excluded from the analysis those plural responses which were clearly the result of a simple error in the root morpheme. These included changes in a single feature such as the devoicing of the /z/ in PHASES to give FACES, insertion of an additional phoneme such as PHRASES instead of PHASES (an error made by 3 subjects) and the homophone error of PEARS for PAIRS (4 subjects). To include such errors would be to bias the results in favour of phonological coding.

We were left with 154 errors to examine from a total of 774 lists. There were 31 changes involving the context set; these are special and will be discussed further below. There remain a total of 123 errors whose distribution is given in Table 2. In 84 cases a plural was added to one of the test set nouns, there were 36 plurals on intrusion nouns, and 3 cases of intrusions involving verbs with 3rd person singular endings.

Table 2. *Errors involving a switch of the plural to another word*

		Stimulus list type			
		/əz/	/s/	/z/	Total
Error type	/əz/	29	10	4	43
	/s/	2	23	16	41
	/z/	11	8	20	39
Total		42	41	40	123

It should be recalled that the test lists contained a plural or plurals of one particular class (/s/, /z/, or /əz/). For each class we tested the null hypothesis that the plural intrusions were spread over the same phonological class or one of the different phonological classes with probabilities of 1/3 and 2/3 respectively, using 2-tailed binomial tests with a correction for continuity. This test requires independence of the occurrences that are counted in the columns of the matrix in Table 2. Although this requirement is not completely met it is closely approximated since we have about 40 errors in each column spread over 86 subjects, and any one subject rarely contributes more than one error to the same column. In the exceptional cases there was no evidence of errors due to a single subject being related. The null hypothesis was rejected at the 1% level for the /z/ and /s/ lists and at the 5% level for the /z/ lists. We can be satisfied, then, that this class of errors is biased towards the same phonological class as the stimulus lists. Such biases could not exist unless the plurals concerned were coded phonemically.

Duplication errors

A further class of errors involved those 267 lists where the plural nouns were correctly recalled. On 29 occasions one of the test set nouns was given a plural, there were 5 intrusive plurals and 18 errors involving the context set. Ignoring the latter for the moment we have 34 errors (13%) whose distribution is shown in Table 3. The columns of this matrix reject the null hypothesis only in the case of the /s/ lists ($p < 0.01$). For the case of the /s/ lists, then, we have weak evidence of a response bias induced by the stimulus list. In practice this could be brought about by the reduplication of a phonologically coded plural morpheme in memory. Precise accounts will depend upon the theoretical superstructure one wishes to adopt. The extent to which this is a matter of general experimental set can be estimated by the fact that in answer to the 258 control lists which contained no plurals, only 9 plurals (3.5%) were given in a reply vs the level of 13% for the errors of duplication.

Table 3. *Errors involving the addition of a plural*

		Stimulus list type			
		/əz/	/s/	/z/	Total
Error type	/əz/	4	2	4	10
	/s/	2	11	6	19
	/z/	1	3	1	5
Total		7	16	11	34

We have treated plural errors involving nouns and verbs as equivalent in Table 2. This is legitimate if one is considering a phonological code. The problem is more complex if one is thinking in terms of morphemes since the inflected form, e.g., SINGS, is restricted to the 3rd person singular present. If the inflection is to be associated with number it would thus be considered singular. If errors involving verbs were thought to be morphemic then it would be the [SINGULAR] marker not the [PLURAL] one which could lead to the error (the addition of an *s*). If this were the case then we would expect a higher proportion of inflected verbs in the control lists. That is if the stimulus list included:

NOUN + PLURAL *dogs*

i.e.

VERB *know*

the possible error would lead to:

NOUN *dog*

i.e.

VERB + PLURAL *know.*

With the control lists we would have:

NOUN + SINGULAR *dog*

i.e.

VERB *know*

changing to the error form:

NOUN *dog*

i.e.

VERB + SINGULAR *knows*

(with assumptions about the nature of unmarked number and person). In fact there were no inflected verbs given as responses to the control lists. Errors involving verbs can provide further evidence in favour of low level coding. The context sets with a single plural noun also included a verb. These verbs were chosen to be of the same phonological class as the plural noun, *sit, know* and *teach* being associated with *tops, beds,* and *spaces* respectively. If the plural is coded phonemically then if an error were made it could as well end up with the verb (giving *sits, knows* or *teaches*) as with the test set noun of the same class. In those 258 lists there were 27 errors involving same class nouns and 29 involving verbs. This could not be the case if morphemic or semantic information were carried with the inflection.

Evidence for morphemic code

We have presented the evidence supporting a phonological code. The existence of such evidence does not preclude the possibility of a morphemic code also being involved. The main evidence for a morphemic code is that the

biases in Table 2 are incomplete. In particular a /z/ plural, lost from the stimulus plural reappears equally often as /z/ and /s/. There seem to be two possibilities in respect of such errors; either they arise from morphemic coding (from some of the subjects or all of the subjects some of the time—there are insufficient errors from individual subjects to evaluate these alternatives) or the phonologically coded plural is subject to some kind of transformation as a result, perhaps, of interference or decay. In such a case we must observe that there are 51 such errors of the 123, giving a probability of 0.41 that a detached plural will be altered in form. While we have no base-rate against which to compare this figure it does seem rather high.

Eight of the context sets contained singular nouns which form irregular plurals. These nouns were included in an exploratory way only, as it proved difficult to match them all in frequency of singular and plural forms with other nouns. They were recalled in plural form on 20 occasions, involving the responses of WOMEN (11 times), MEN (3), FEET (2), TEETH (2), and MICE (2). LICE, GEESE and PENCE were never given as responses. Again we have trouble with a base-line against which to evaluate such figures. The responses all involve a change in the vowel alone (in both vowels in the case of WOMEN). Looking at responses to the other words in the stimulus lists we found only four which involved just a vowel change; FAR was given twice for FIRE and TIPS and SET were given once each for TOPS and SIT. A large number of other similar possibilities never occurred such as ROSE–RISE, LIKE–LOOK, HEART–HURT, FOOT–FIT, SHIP–SHOP, WALL–WOOL. We take the 20 irregular plurals, then, as weak evidence in favour of some morpheme coding of the plurals. More detailed discussion of the implications of this kind of error would be contingent on the way chosen to describe such plurals (see Palmer, 1971, p. 112 ff. for an account of some of the problems).

It seems, then, that we have evidence of at least two forms of coding of the plurals. The position with respect to the root nouns is equally complex. The existence of homophonic substitutions might be taken as evidence in favour of the essentially phonological code of the material being remembered. Thus WRITE was variously reproduced as RIGHT (4 times) and WRIGHT (twice). However, we may view this data from the other side and point out that in about 50 cases WRITE was reproduced correctly. Such accuracy might be claimed to constitute evidence in favour of a visual or a semantic code at least supplementing the phonological code in the majority of cases. The other examples of phonological errors were COURT which was once produced as CAUGHT and THERE which was once produced as THEIR.

In conclusion we claim to have shown in a situation where subjects are visually presented with a list of words for immediate serial recall, that plurals are coded separately from the root nouns they modify. There is evidence

from the distribution of errors showing that the plurals are primarily coded in a phonological form but some other evidence suggests the presence of a morphemic code some of the time. It is possible that the two-kinds of error occur at different stages of the task (e.g., storage vs retrieval): equally it is possible that information is liable to be coded in many forms at the same time.

References

Glanzer, M. and Razel, M. (1974) The size of the unit in short-term storage *J. verb. Learn. verb. Behav., 13,* 114–131.

Kucera, H. and Francis, W. N. (1967) *Computational Analysis of Present-day American English.* Rhode Island, Brown University Press.

Morton, J. (1968) Grammar and computation in language behaviour. In J. C. Catford (ed.), *Studies in Language and Language Behaviour* C.R. L.L.B. Progress Report, 6, University of Michigan.

Morton, J. (1969) Interaction of information in word recognition. *Psychol. Rev., 76,* (2), 165–178.

Morton, J. (1970) A functional model for memory. In: Norman, D. A. (ed.), *Models of Human Memory.* New York, Academic Press.

Murrell, G. A. and Morton, J. (1974) Word recognition and morphemic structure. *J. exper. Psychol., 102,* (6) 963–968.

Osgood, C. E. and Hoosain, R. (1974) Salience of the word as a unit in the perception of language, *Percep. Psychophys., 15,* 168–192.

Palmer, F. R. (1971) *Grammar,* Harmondsworth, Penguin Books.

Taft, M. and Forster, K. I. (1975) Lexical storage and retrieval of prefixed words. *J. verb. Learn. verb. Behav., 14,* 638–647.

APPENDIX

Testword Sets

noise	rose	chance	edge	bridge	price
fruit	lake	heart	ship	park	test
queen	king	club	bar	team	game
horse	house	case	voice	axe	inch
unit	fact	point	court	lamp	guest
wall	home	boy	table	lawn	yard

Context Sets

from	gay	there	tops	beds	spaces
warm	far	queen	dark	man	feet
sing	climb	write	sit	know	teach
rocks	trains	phases	cups	pairs	fences
aunts	pens	ranches	books	plans	pieces
mouse	woman	louse	goose	penny	tooth

Résumé

On a demandé à des sujets de répéter des listes de 6 mots présentés visuellement. Parmi ces mots se trouvaient un ou deux noms au pluriel. On trouve un nombre important d'erreurs qui se traduisent par le détachement de la forme plurielle de la racine originale. On considère cela comme une preuve en faveur d'un codage fondé sur le morphème plutôt qu'un codage fondé sur le mot comme unité. Dans une proportion significativement élevée ces erreurs maintiennent la forme du pluriel (/s/, /z/ ou /əz/) dont on pense qu'elle prouve la priorité du codage phonologique du morphème pluriel. Cependant il est possible que le pluriel soit parfois codé morphonémiquement c'est à dire de façon à ce que les diverses terminaisons plurielles ne soient pas distinguables.