

Distributed versus Localist Representations: Evidence from a Study of Item Consistency in a Case of Classical Anomia

Matthew A. Lambon Ralph

MRC Cognition and Brain Sciences Unit, Cambridge, United Kingdom

Models of word production and comprehension can be split into two broad classes: localist and distributed. In localist architectures each word within the lexicon is represented by a single unit. The distributed approach, on the other hand, encodes each lexical item as a pattern of activation across a set of shared units. If we assume that the localist representations are more than a convenient shorthand for distributed representations at the neuroanatomical level, it should be possible to find patients who, after brain injury, have lost specific words from their premorbid vocabulary.

Following a closed head injury, JS had severe word-finding difficulties with no measurable semantic impairment nor did he make phonological errors in naming. Cueing with an initial phoneme proved relatively ineffective. JS showed a high degree of item consistency across three administrations of two tests of naming to confrontation. This consistency could not be predicted from a linear combination of psycholinguistic variables but the distribution fitted a stochastic model in which it is assumed that a proportion of items have become consistently unavailable.

Further evidence is presented which suggests that these items are not, in fact, lost but rather have a very low probability of retrieval. Given phonemic cueing of sufficient length, or delayed repetition priming from a written word, the consistently unnamed items were produced by JS. Additional data is reported which seems to support a distributed model of speech production. JS's naming accuracy for one set of pictures was found to predict his performance on a second set of items only when the names of the pictures were both semantically and phonologically related (e.g.,

This work was completed while the author was supported by the Engineering and Physical Science Research Council. I am particularly grateful to Jo Roberts, who originally referred JS to me. I thank Karalyn Patterson, Andy Ellis, David Howard, and David Plaut for their useful comments on earlier drafts of this paper and David Howard for providing me with valuable data on the Hundred Picture Naming Test. Finally and most importantly, I am indebted to JS for his patience and willingness to complete hours of testing.

E-mail: matt.lambon-ralph@mrc-apu.cam.ac.uk.

Address correspondence and reprint requests to Matthew A. Lambon Ralph, MRC Cognition and Brain Sciences Unit, 15 Chaucer Road, Cambridge CB2 2EF, United Kingdom. Fax: 01223.359062.

cat-rat). There was no association for pairs of pictures if they were only semantically (e.g., cat-dog) or phonologically related (e.g., cat-cap).

It is argued that JS's data are best described in terms of a graded, non-linear, distributed model of speech production. © 1998 Academic Press

INTRODUCTION

Models of word recognition and production can be split into two broad classes. One collection of models contains representations in a localist form, i.e., they contain single entities or nodes each of which corresponds directly to one, and only one, representation. Distributed models, on the other hand, do not use this one-to-one correspondence, but rather code each representation as a pattern of activation across a number of elements; each element can be utilized as a component of many different representations. The logogen model of Morton (1985) and the interactive activation model of Dell and colleagues (Dell, 1986; Dell & O'Seaghdha, 1992) both contain a single layer of units in which each unit corresponds to a single lexical item. Similarly, two-stage models of speech production (Butterworth, 1989; Levelt, 1989, 1992) contain two sequential layers of localist representations. Distributed models (e.g. Plaut & Shallice, 1993; the "triangle" model: Seidenberg & McClelland, 1989; Plaut, Seidenberg, McClelland, & Patterson, 1996), on the other hand, treat speech production as a process in which both semantic and phonological representations are coded across a number of nodes. Although they employ a set of hidden units between the semantic and phonological layers, this cannot be thought of as a lexicon in the traditional sense because, for different words, considerably overlapping sets of hidden units participate in activating phonology from semantics.

Localist and distributed representations predict different patterns of impairment in aphasic production. If representations are stored (in the brain) in a localist form¹ then it should be possible to identify patients whose anomia is best conceptualized as resulting from the loss of specific lexical entries. If a selection of words has been removed from patients' premorbid vocabulary, it should be impossible for these items to be retrieved under any circumstances. Such patients should *never* produce the picture names even if the tests are accompanied by procedures designed to facilitate production such as phonemic cueing or priming.

Distributed models predict that particular single items cannot be entirely lost while leaving the remaining vocabulary undamaged. These models are, in fact, resistant to small amounts of damage while greater degrees of impairment lead to a proportional increase in dysfunction across the entire vocabulary, a property known as *graceful degradation* (McClelland, Rumelhart &

¹ Not all researchers assume that localist units within a computational model translate directly into localist, neuroanatomically based representations (see "Discussion") but some authors do (see Howard, 1995; Page, submitted).

Hinton, 1986). Distributed models of speech production would predict that some items may be consistently unavailable if they are relatively impoverished (perhaps due to representation across a small number of nodes and low weighted connections between sequential layers, a disadvantage which might be exacerbated by brain injury). Specifically, low frequency, late-acquired words relating to unfamiliar, abstract concepts will have such a low probability of production that in practice they will be consistently unavailable and might appear to be lost (see Coltheart & Funnell, 1987, for a rehearsal of a similar argument in respect of orthographic representations). Although these words may be relatively hard to retrieve, it should be possible to facilitate production with cueing or priming, at least for a proportion of the consistently unavailable items.

Howard (1995) presented a single case study of an anomic patient EE. Following a closed head injury, EE demonstrated very consistent item-by-item performance across repeated administrations of the Hundred Picture Naming Test (Howard & Franklin, 1988). This severe but consistent anomia presented with no apparent central semantic impairment nor any pure phonological disturbance. With regard to comprehension, EE performed within the normal range on both the picture and spoken-word versions of the Pyramids and Palm Trees Test (Howard & Patterson, 1992) and scored normally on the spoken version of the Peabody Picture Vocabulary Test (Dunn, 1965). His score on the written version of this test was lower than the normal range although this was presumed to be due to an additional impairment to input orthographic processes and not reflective of a mild central semantic impairment: in Howard & Franklin's (1987) terms an impairment to the visual input lexicon. Despite apparently intact semantics, EE was severely impaired at picture naming on standardized tests (no correct responses on the Graded Naming Test: Warrington & McKenna, 1983; 15/60 and 13/60 correct on two administrations of the Boston Naming Test: Kaplan, Goodglass, & Weintraub, 1976).

Howard argued that EE's pattern of impairment was best thought of as a loss of specific lexical items. He demonstrated that 30/100 items from the Hundred Picture Naming Test were consistently unavailable (across three separate administrations). The alternative hypothesis, that the words were consistently unavailable because they were relatively hard to retrieve, was rejected in favor of the localist theory on the basis of two additional pieces of evidence. First, statistical models based on a linear combination of 11 psycholinguistic variables provided a very poor fit to the observed consistency distribution while stochastic models, which assume that a proportion of items are consistently unavailable, predicted a pattern of performance very close to that of EE's. Second, EE's naming was only marginally improved by initial phonemic cues or extra time.

Several questions arising from this study remain unanswered. EE's naming performance was not perfectly consistent. Thirty percent of the target

names were unavailable to EE on each of the three administrations and 50% were consistently correct; but another 20% were inconsistently available. What happened to these latter items? Was EE's naming damaged in two separate ways, one which erased a proportion of lexical items and another which left a selection of words difficult, but not impossible, to retrieve? If this was the case, why were the inconsistent items in a medium range of rated word familiarity (mean = 543) falling between the "lost" items (mean = 503) and the consistently named items (mean = 574)? This might suggest that the putatively lost words were not in fact "missing" but rather were consistently unavailable because their probability of retrieval was so low (if this probability is based on word familiarity in some way).

Presented below are data from another classical anomic patient, JS. The analysis of his data is split into four separate sections. The first gives background neuropsychological data. The second shows that JS's naming performance on the Hundred Picture Naming Test was very consistent and that this consistency was not predicted by any linear regression models but did fit with stochastic models which assume that a proportion of items are "lost" (Faglioni & Botti, 1993). However, if the consistently unnamed pictures were accompanied by a phonemic cue of sufficient length, then JS was able to retrieve all the target names. In the third section, JS's item consistency was replicated across three administrations of the Boston Naming Test. Again the consistency was much greater than that predicted by a linear regression model, and consequently, it could be concluded that the consistently unavailable items had been removed from JS's vocabulary. If JS was required to read aloud the target names a short while before he was presented with the picture, he was able to name a significantly greater number of pictures including a proportion of the consistently unavailable items. The final section demonstrates that JS's unnamed items did relate to the naming accuracy of other pictures, not for phonologically *or* semantically related pairs but for those items that were both phonologically and semantically related.

CASE REPORT

JS was a 41-year-old businessman. In May 1995, he suffered a head injury following a bicycle accident. This left him unconscious for a period of 2 weeks during which he underwent neurosurgery to drain a hematoma. In June of the same year he was hit on the head during an assault. It is unclear whether this second incident resulted in any additional brain damage. A CT scan carried out 1 month later (July) revealed a large low density area in the left temporoparietal region. A follow-up scan (performed in November) showed a wedge-shaped low density area in the left temporal lobe and a slight dilation of the temporal horn of the left lateral ventricle. Testing began in March of the following year, and the majority of results reported here were collected in August 1996, some 15 months post the first injury. JS

TABLE 1
JS's Performance across a Variety of Semantic Tests

Test	JS	Normal range
PALPA ^a Written Word-to-Picture Matching	39/40	35–40
PALPA ^a Spoken Word-to-Picture Matching	39/40	35–40
BORB ^b Picture Association Match	30/30	21–30
Pyramids and Palm Trees Test ^c (all picture version)	51/52	49–52
Shallice and McGill Spoken Word-to-Picture Matching ^d		
Concrete	30/30	27–30 ^e
Abstract	27/30	21–29 ^e
Emotion	15/15	11–15 ^e
PALPA ^a Written Word Association		
Concrete	15/15	$m = 13.4$ (1.26)
Abstract	12/15	$m = 12.3$ (1.82)
PALPA ^a Written Synonym Judgement		
High imageability	30/30	Not available
Low imageability	29/30	Not available

^a PALPA: Psycholinguistic Assessments of Language Processing in Aphasia (Kay, Lesser & Coltheart 1992).

^b BORB: Birmingham Object Recognition Battery (Riddoch & Humphreys, 1992).

^c Pyramids and Palm Trees Test (Howard & Patterson, 1992).

^d Shallice and McGill (unpublished test).

^e Norms collected in York from 10 undergraduates (ages 19–32).

presented with fluent speech with occasional word-finding difficulties. Initial testing revealed no apparent comprehension difficulties, and no problems with single word repetition or reading aloud apart from a small number of visual paralexia. Further data on JS's naming is reported elsewhere (Lambon Ralph, Sage, & Robertson, submitted).

Background Neuropsychological Testing

JS performed a variety of tests requiring semantic processing with extreme ease and efficiency. As shown in Table 1, across these semantic tests JS scored either perfectly or at the high end of the normal range. These results substantiate the strong clinical impression of no apparent comprehension deficit.

Although able to give detailed and precise information about target pictures, JS was severely anomie. On the Boston Naming Test (Kaplan, Goodglass, & Weintraub, 1976) he correctly named 16/60 target pictures (normal range 42–60) with 3 circumlocutions, 3 semantic errors (subsequently rejected), and 38 non responses. On the Graded Naming Test (Warrington & McKenna, 1983) he named only 4/30 items correctly (normal subjects mean = 22.5 items, s.d. = 4.3); here JS's errors were 2 circumlocutions and 24 nonresponses. This pattern of performance was repeated across a large num-

ber of picture naming tests. Phonological errors were never produced and his most common error type was to make no response (with occasional circumlocutions where the time given for each picture allowed it). JS did make a small number of semantic errors although he rejected them (again suggesting good comprehension). In this respect, JS is another patient who produces semantic errors in naming without a central semantic impairment (Caramazza & Hillis, 1990).

Cueing with initial phonemes had a minimal effect. On the Graded Naming Test only six additional items were elicited following the phonemic cue while another led to a semantic error (*radius* + /r/ → ‘rhombus’). On a larger set of Snodgrass and Vanderwart (1980) pictures (presented twice, once with no cue and once with the initial phoneme, counterbalanced across sessions) the phonemic cue increased JS’s performance from 87/140 to 116/140, a change of borderline significance (McNemar $\chi^2 = 3.23, p = .07$).

Naming Performance on the Hundred Picture Naming Test (Howard & Franklin, 1988)

It appeared that JS regularly failed to name a number of specific target items, even if the picture (albeit in a different form) was repeated in a subsequent test in the same session. For example, having been unable to name a picture of (a wedge of) cheese on one occasion, he was asked to produce it again in a later test. He was unable to name it the second time despite having been told the correct label on the first occasion. Interestingly this same problem presented in real life as JS had considerable difficulties when ordering a cheese sandwich because he could “never remember the name.” This anecdotal consistency was more formally assessed using the Hundred Picture Naming Test.

Method

The Hundred Picture Naming Test was presented three times, 1 week apart, with the items occurring in the same order on each occasion and with a 5-s time limit per item. On the first two occasions no cues were given; on the last occasion, phonemic cues were given for unnamed items. The precise method and results of this phonemic cueing are presented below.

Results

JS correctly named 51 items on the first occasion (38 nonresponse errors and 11 semantic errors including 5 superordinate responses). On the second occasion he named 52 items (40 nonresponses and 8 semantic errors including 4 superordinates). On the last occasion he named 53 pictures (39 nonresponses, 1 circumlocution and 7 semantic errors including 1 superordinate). Comparing across sessions, 35 items were named correctly on all three occasions, 33 items were never named and 32 items were inconsistently named (13 once and 19 twice). Contingency coefficients calculated between pairs

TABLE 2

Simple Correlations between JS's Naming Success on the Hundred Picture Naming Test and the Psycholinguistic Properties of the Items, Together with Means, Standard Deviations, and Ranges

Variable	<i>r</i>	<i>p</i>	Mean	SD	Range
Log Kucera and Francis frequency	.314	.001	1.36	.6	0–2.65
Rated age-of-acquisition	–.298	.003	261	60.9	150–436
Rated visual complexity	.078	n.s.	3.49	1.2	1.6–6.1
Rated concreteness	.175	n.s.	593	33.1	482–638
Rated word familiarity	.294	.003	546	59.0	386–645
Rated imageability	.351	<.001	590	31.1	483–639
Number of letters	.005	n.s.	4.7	1.3	3–11
Number of phonemes	.016	n.s.	3.6	1.0	2–8
Number of syllables	.106	n.s.	1.3	.5	1–4
Rated operativity	.058	n.s.	4.6	1.0	2.1–6.2

of administrations also showed a high degree of consistency ($C = .527$ for administrations 1 and 2, maximum possible value = $.700$; $C = .489$ for 1 and 3, maximum possible value = $.693$; $C = .474$ for 2 and 3, maximum possible value = $.700$).

Was JS's consistency greater than that expected by a linear combination of psycholinguistic variables? (Note the contingency coefficient is assessed against a chance level.) Ten psycholinguistic variables (see Table 2) were included in the analyses: log transformed written word frequency (Kucera and Francis, 1967); rated age-of-acquisition; familiarity; concreteness; imageability (Coltheart, 1981), the numbers of letters, phonemes, and syllables; rated operativity; and visual complexity (Howard, personal communication). JS's total score (summed across the three administrations for each item) significantly correlated with 4 of these 10 variables: log frequency, rated age-of-acquisition, imageability, and word familiarity.

These intercorrelated variables were tested for independent effects on JS's naming using a simultaneous linear regression. When all ten variables were entered into the regression, only imageability ($t = 2.41$, $p = .02$) and log frequency ($t = 2.22$, $p = .03$) were found to have significant independent effects. Rated age-of-acquisition and word familiarity failed to reach significance. This picture remained unchanged when only one variable of each type (e.g., one variable for length) was entered into the regression. For example, imageability and log frequency remained as the only significant predictors when four other factors (age-of-acquisition, visual complexity, number of syllables, and operativity) were included in the equation. Rated familiarity reached borderline significance only if it was substituted for log frequency ($t = 1.78$, $p = .08$).

The linear regressions can provide a series of null hypotheses against which JS's naming consistency can be gauged. The linear regression equa-

tion was used to estimate the probability of naming each picture. This estimate was used in the binomial expansion to calculate the probability of correctly naming each item on zero, one, two, or three occasions. These values were summed across the 100 items and the resulting predicted consistency distributions were compared to that shown by JS (NB this is the same method used previously for patient EE: Howard, 1995). The results are shown in Table 3.

Various single variables or combinations of variables were used in the linear regression models. Each was an improvement on the distribution predicted by chance alone. Even with the best linear equation (containing all 10 variables), however, a large discrepancy between the predicted and observed consistency distribution still remained. In this case the value of the χ^2 statistic is very high ($p < 10^{-7}$).

If JS's performance was due to a permanent loss of certain lexical items then a stochastic Markov chain model (Faglioni & Botti, 1993) should yield a much better fit to the data. This three parameter model provided an excellent match to JS's data (sufficiency, $\chi^2_{(4)} = 1.16$, n.s.). The estimated proportion of consistently unavailable items was 0.300 (95% confidence interval of 0.198–0.402); the proportion that were automatically retrieved was estimated to be 0.368 (95% confidence interval of 0.178–0.557); and the probability of the remainder being retrieved was estimated to be 0.594 (95% confidence interval of 0.424–0.764).

Summary

The results of these analyses show that JS's consistency was significantly greater than that expected on the basis of a linear combination of 10 psycholinguistic variables. Even when compared to the model that provides the best fit to JS's naming accuracy, there is still a large difference between the predicted distribution and that shown by JS. The shape of each is quite different; the predicted distribution is "hill shaped" (maximal for medium values, i.e., inconsistent naming) rather than being "U shaped" like that for JS (see Table 3). On the other hand, a stochastic model provides a close fit to the data which predicts that 30% of the items were consistently, i.e. permanently, unavailable to JS. Even if we take the lower end of the 95% confidence interval, the model predicts that at least 20% of the items were consistently unavailable.

Cueing the "Lost" Items

A proportion of items was consistently unavailable to JS. If these were permanently lost rather than very hard to retrieve then phonemic cues should be ineffective. Previous testing of JS's naming (see the above section on Background Neuropsychology Testing) had revealed that initial phonemic cues provided limited assistance. For the majority of targets, however, there

are many more phonemes in the name that could be incorporated into a cue. If we take a model of speech production in which a proportion of lexical items have been erased, then no amount of phonemic cueing should prove successful. To assess whether specific lexical items are missing on the basis of data from phonemic cueing, one should provide cues up to the phonological uniqueness point of the target word.²

To test the efficacy of additional phonological information, the last administration of the Hundred Picture Naming Test was given with phonemic cueing. When JS failed to name a picture in 5 s, phonemic cues of increasing length were given. For each unnamed item, the cue was repeated with an additional phoneme every 5 s until JS correctly named the target or the cue contained the entire picture name. This procedure is similar to gating (Tyler, 1992) in the presence of a picture of the target (Graham, Patterson & Hodges, 1995; Wingfield, Goodglass, & Smith, 1990).

The length of the cue required to elicit the target label for the consistently unnamed items is shown in Table 4. Out of the 33 consistently unnamed items only 1 (wool) failed to be cued before the entire label had been given, and it turned out that JS had been trying to retrieve a related verb (knitting) which is appropriate for the picture. Ten out of the 32 items were cued with the initial phoneme, the remainder requiring two ($N = 15$) or three phonemes ($N = 7$).

As many of the target names were CVC syllables, it could be that two phonemes constitute the uniqueness point of many items, thus narrowing the phonological cohort to an N of 1, i.e., the target word itself. Successful cueing of this kind would be a less- although not an uninteresting result, but further analysis proved it not to be the case. The MRC Psycholinguistic Database (Quinlan, 1992) was used to estimate the size of the phonological cohort for each successful phonemic cue given. The results are shown in Table 4. The cohort size was calculated to be the number of words that shared the same initial phonemic string as the successful cue. The mean cohort size was found to be 490, representing 0.5% of the corpus size. The smallest cohorts were for the items globe and pepper: after a 3-phoneme cue in each case, the cue specified a cohort of eight words. When the phonological cohort was analyzed for its frequency distribution, the target word was found to be in the upper half. The target frequencies were significantly greater than the median cohort frequency ($t = 3.49$, $p = .001$) or, to put it another way, 24/32 items had a target frequency greater than or equal to the median cohort

² On a localist model of speech production, it should perhaps be possible to give the entire word as a cue without the patient retrieving the label because the entire word-form has been erased. However, if we assume that there are separate input and output representations, then the entire label as a cue would reduce the "input" cohort to only one item and the patient would be able "name" the picture by repeating the cue. In fact, for semantically intact patients like JS, the word will also match their semantic representation derived from the picture.

frequency. It is possible that this represents a contributing factor to successful retrieval—the phonemic cues not only reduce the size of the target cohort, they also typically make the target item one its most frequent members.³

Naming Performance on the Boston Naming Test (Kaplan et al., 1976)

The items from the Boston Naming Test were used to investigate JS's performance in order to replicate the high item consistency found on the Hundred Picture Naming Test, which could not be predicted by a linear combination of psycholinguistic factors and to demonstrate that JS could name the consistently unavailable items with assistance other than phonemic cueing.

Method

In order to assess JS's item consistency, the Boston Naming Test was administered three times (without any cues). In addition, JS was asked to name the pictures after two forms of repetition priming from a written word. The delay between prime (reading the word) and target was changed for each condition. In the short-delay condition JS read the prime word, named an unrelated picture, read another word, and then attempted to name the target picture (i.e., there were two unrelated stimuli between prime and target). In the long-delay condition, JS was asked to read a list of the target names, he was then engaged in another activity for 15 min after which the naming test was administered (without any primes or cues). There was at least 1 week between each of the five testing sessions. In order to ensure that any improvement in naming accuracy was not due to practice, the following order was adopted: (1) plain administration, (2) plain administration, (3) short-delay cueing, (4) long-delay cueing, and (5) plain administration. If there was any practice effect, JS's performance should have been better in the last testing session than in either of the first two.

Results

JS's accuracy was very consistent across the testing sessions (without any priming). He correctly named 24/60 on the first occasion, 25/60 on the second and 22/60 in the last administration. Consistency coefficients calculated between pairs of administrations were very high (between sessions 1 and 2: 0.61; sessions 1 and 3: 0.58; sessions 2 and 3: 0.53).

JS's consistency distribution is shown in Table 5. There were 30 items (50%) which he never named, 6 (12%) which he named once, 7 (14%) which he named twice, and 17 (34%) which he named on each of the three occasions. The observed distribution was assessed against a linear combination of two psycholinguistic variables (using the same method as described above): (log) frequency per million drawn from the Celex database (1993)

³ Given the results of the section entitled "Do Unnamed Items Predict Naming Failure on Other Pictures?" it seems probable that the number of semantically related items in the phonological cohort may also be a factor. It is most likely that the phonemic cue is successful because it reduces the number of phonologically and semantically related targets as well as making many target labels above average frequency in their cohorts.

TABLE 4
 Statistics on the Consistently Unnamed Pictures and Their Successful Phonemic Cue

Picture	K-F frequency ^a	Length of cue (/phonemes)	Cue's phon' cohort ^b	% of corpus ^b	No. of cohort in K-F ^a corpus	Median K-F frequency ^a
Arch	13	1	198	.20	87	4
Bag	42	2	203	.21	84	5
Bow	15	1	1631	1.66	764	4
Bridge	98	3	10	.01	8	7
Brush	44	2	243	.25	119	3
Cheese	9	1	235	.24	108	5
Cotton	38	3	10	.01	3	19
Crook	3	2	323	.33	133	3
Dress	67	3	11	.01	5	6
Duck	9	1	2153	2.18	901	4
Flag	16	3	52	.05	19	1
Fork	14	2	143	.15	62	6
Frog	1	2	200	.20	84	5.5
Globe	13	3	8	.01	2	14.5
Goal	60	2	27	.03	13	2

Hedge	2	1	1129	1.15	465	4
Hoof	2	2	17	.02	11	8
Horn	31	1	1129	1.15	465	4
Hose	9	2	53	.05	28	9.5
Mayor	38	1	1624	1.65	734	4
Mermaid	1	2	37	.04	19	7
Onion	15	2	447	.45	167	2
Palm	22	2	95	.10	49	6
Pepper	13	3	8	.01	2	8
Post	84	1	2618	2.66	1118	97
Rake	11	2	87	.09	48	5
Skirt	21	2	306	.31	121	3
Step	131	2	487	.49	251	5
Tent	20	2	228	.23	90	7.5
Thermometer	10	1	216	.22	99	5
Toe	9	1	1728	1.75	711	4
Worm	4	2	40	.04	25	9
Mean	27.0	1.88	490.5	.5	212.3	8.7

^a Kucera and Francis Corpus (1967).

^b The number of words contained in the (successful) phonemic cue's phonological cohort taken from MRC Psycholinguistic Database (Quinlan, 1992).

TABLE 5

JS's Consistency across Trials on the Boston Naming Test Assessed against Two Statistical Models and the Number of Items Retrieved with the Assistance of an Identity Prime

No. of times each picture is named correctly	JS's consistency distribution/ 60 pictures	Consistency distribution predicted by		No. of items named after priming with the target label (short-delay condition)
		Chance	Phoneme length and log frequency	
0/3	30	13.3	17.6	11/30
1/3	6	26.0	21.0	3/6
2/3	7	17.0	14.2	5/7
3/3	17	3.7	7.2	15/17
Difference between observed and predicted distributions	$\chi^2_{(3)}$ p	90.3 $\ll .001$	36.4 $\ll .001$	

and phoneme length (the 10 factors included in the analysis of data from the Hundred Picture Naming Test were not available for the Boston Naming Test). The predicted consistency distribution is shown in Table 5. Again, there was a large discrepancy between the predicted and observed patterns ($\chi^2 = 36.4, p < 10^{-7}$).

Long-delay priming had no beneficial effect on JS's naming (24/60 correct) but in the short-delay condition his accuracy was improved to 34/60 (this score was significantly better than his naming in sessions 1 or 3: McNemar $\chi^2 = 4.5, p = .03$; McNemar $\chi^2 = 7.6, p < .01$; although the difference failed to reach significance for session 2: McNemar $\chi^2 = 2.8, p = .09$). The relationship between the number of items primed and the consistency distribution is shown in Table 5. Although priming was more effective for pictures that JS had named on at least one occasion than those he had never named (23/30 vs. 11/30: $\chi^2 = 8.2, p = .004$), even a proportion of the consistently unavailable items were named if they had been preceded by the written label. This outcome is not predicted if the consistently unnamed pictures are considered to have been deleted from the output lexicon.

Do Unnamed Items Predict Naming Failure on Other Pictures?

Distributed and localist models provide different hypotheses with regard to this question. It is possible that in a distributed system naming accuracy for one picture may predict success on other semantically or phonologically related items because the representations are encoded across the same set of units (see Allport, 1985). This does not follow from localist models: if a word cannot be named because it has been "removed" from the lexicon, it will not prevent the remaining words from reaching threshold. In his analysis of patient EE, Howard (1995) failed to find an association between naming

accuracy for phonologically related items and only a borderline correlation for semantically related items (which did not remain when familiarity was partialled out).

Patients like JS do not have any documented abnormality of semantic representations, nor do they make phonological errors in naming, repetition, or reading, suggesting that phonological representations are likewise unperurbed. In distributed models of production, their impairment probably lies somewhere on the link between semantics and phonology (an idea used by other authors to explain the profound anomia observed in some cases of semantic dementia: Graham, et al., 1995). For patients with an impairment between semantics and phonology, the probable association for naming accuracy is not between semantically or phonologically related pictures but rather between those items that are both semantically and phonologically related. Consider the following words: *cat*, *rat*, *mat*, and *dog*. Like the majority of semantically related words *cat* and *dog* have unrelated phonological word-forms. Thus the transform from semantics to phonology required for these two items is very different (two similar semantic representations are transformed into unrelated phonological forms). The same is true for *cat* and *mat* although for an alternative reason: here, phonology but not meaning is similar (a different transformation is required for each because they start from unrelated semantic representations). For *cat* and *rat* the position is different. The two items are both semantically and phonologically related and, in a distributed system, there would be considerable overlap of features at both levels of the computation (as the two start and end with similar representations the transformation required for each will be similar—brain damage that affects the transformation of *cat* from meaning to sound will also impair performance for *rat*). Therefore one should be most likely to see an association for this latter pair.⁴ For other types of patients, the association may change with the locus of their impairment. For instance, patients with a category-specific semantic impairment, by their very nature, show an association between semantically related pictures. This follows from a distributed model where, for example, the naming of animals relies disproportionately on a collection of features. If a subset of these attributes is degraded then the naming of all concepts which rely, in part, on the missing features will be compromised.

Consider the same words in a localist architecture. If the lexical node for *cat* has been removed then *rat*, *dog*, and *mat* should still reach threshold

⁴ The position described here is still a simplification. The naming accuracy for any specific item will depend, at least in part, on the richness of its semantic representation, how "typical" the phonology is (in the sense of high frequency phonemes, high probability transitions between successive phonemes, etc.), and how often we transform between the two representations. These factors will alter the expected association between semantically and phonologically related items. Consider the items "sow" and "cow." If the patient did not name "sow" we would not be surprised if the patient nevertheless correctly named "cow."

(although *rat* will no longer receive the same degree of feedback activation previously provided by the *cat* lexical node: see Dell & O'Seaghda, 1992 for details).

The prediction arising from consideration of a distributed architecture was tested using the target items from a spoken-word-to-picture matching test (Franklin, Turner & Ellis, 1992). In this test there are three sets of picture pairs. In one set, the pictures have phonologically related names, in another the pairs are semantically related, and in the third set the pictures are semantically *and* phonologically related. JS was asked to name all the pictures from the battery, presented in a pseudo-random order such that related pairs did not occur in succession.

JS named 50/92 (54%) of the phonologically related set, 54/90 (60%) of the semantically related items, and 38/82 (46%) of the semantically and phonologically related pictures. For each set, JS's performance on one member of each pair was compared to his naming accuracy for the other (e.g., did JS's naming accuracy for a picture like *cat* correlate with his performance on an item like *rat*?) The naming results for phonologically or semantically related pairs did not reveal a significant association ($\chi^2 = 2.33, p = .13$; $\chi^2 = 1.29, p = .26$; respectively). A small but significant association was found for pairs of pictures that were both semantically and phonologically related ($\chi^2 = 4.36, p = .04$). This result is impressive given that not all the target items are especially close in their semantic relationship (e.g., log and frog, snake and snail, ant and plant).

Howard (1995) found a significant correlation between semantically related items which failed to emerge when rated word familiarity was partialled out. This possibility was tested for the association between semantically and phonologically related pairs. As rated familiarity was not available for the test pictures, the Celex database (1993) was used to provide frequency counts for the picture labels. Using the (log transformed) combined frequency count, no correlation was found between the frequencies of the target pairs ($r = .01, n.s.$). Likewise for (log transformed) spoken frequency, no correlation was found ($r = .02, n.s.$). When either of these frequency counts was partialled out, the raw correlation found between the naming accuracy of semantically and phonologically related items ($r = .33, p = .02$) remained (removing log combined frequency: $r = .36, p = .01$; removing log spoken frequency: $r = .30, p = .03$).

DISCUSSION

This paper presents data from a classical anomic patient. JS showed no evidence for a comprehension deficit nor did he make any phonological errors in naming, reading, or repetition. Across three administrations of two separate tests of naming to confrontation, JS exhibited substantial and significant levels of item consistency. This consistency could not be predicted

by a linear combination of psycholinguistic variables. His naming was not significantly improved by single phoneme cues and failing to name a picture did not predict the accuracy on either phonologically *or* semantically related items.

From this evidence it could be concluded that JS was another patient for whom a proportion of lexical items had been “lost” and that this pattern of performance is incompatible with models of speech production in which word representations correspond to activation over a distributed set of elements (e.g., Plaut & Shallice, 1993; the “triangle” model: Seidenberg & McClelland, 1989; Plaut, Seidenberg, McClelland, & Patterson, 1996). Furthermore, it could be claimed that the pattern shown by JS demands a model which adopts localist representations, such as the logogen model (Morton, 1985), the IAC model of Dell and colleagues (Dell, 1986; Dell & O’ Seaghdha, 1992), or two-stage models (Butterworth, 1989; Levelt, 1989, 1992).

In the framework of a localist model it seems entirely possible that a random selection of lexical items could be lost. There was, however, a clear pattern in JS’s data. For the Hundred Picture Naming Test, the familiarity of JS’s inconsistently named items fell between that of the consistently named and unnamed items (consistently unnamed: mean word familiarity = 524.5; inconsistent items; 545.5; consistently named: 567.9). On the Boston Naming Test, the frequency of the inconsistently named items fell between that of the consistently named and unnamed items (consistently unnamed: mean word frequency (log) = 0.06; inconsistent items: 0.20; consistently named: 0.49). These two observations would seem to point toward a theory in which naming accuracy is graded in some way. If true, the “missing” words are not lost but have such a low probability of retrieval that, in normal circumstances, it is unlikely that the patient will ever be able to name them. This “graded” hypothesis provides a single explanation for the word-finding difficulties present for both consistently unnamed and inconsistently named pictures—the inconsistent items fall at a point on the continuum where name retrieval in the damaged system is just possible whereas the consistent items are so far below this “threshold” that the patient cannot gain access to the picture names. If one adopts the hypothesis in which a set of words has been erased from the lexicon, then there needs to be two explanations for the patient’s anomia—one for the consistently unnamed items (lost) and another for the inconsistently named items. This hypothesis may also give an explanation for the graded efficacy observed in the priming study: JS was more likely to name a picture after reading the written name if it was (at least) an inconsistently named item than if it was a consistently unavailable label. One can imagine that the prime acts by temporarily reducing the “threshold” noted above and consequently a greater number of items will be named. This effect will be greatest for the inconsistently named items because they fall at a higher point on the continuum than the consistently unnamed pictures.

The graded hypothesis predicts that given the right circumstances it should be possible to aid the retrieval of even the consistently unavailable names by some form of cueing or priming. For JS, all the consistently unnamed pictures from the Hundred Picture Naming Test were subsequently named when JS was given a sufficient phonemic cue, normally two or three phonemes. Even at this point in the target word's phonological sequence, at least several other words remained in the phonological cohort (a minimum of 7 and, on average, some 490 words). In addition, JS was able to name a significantly greater number of pictures from the Boston Naming Test (including some of the consistently unavailable items) when he had recently read aloud the name of the target picture.

How similar was JS to EE (Howard, 1995)? There is clearly quite a degree of overlap between the two profiles. Both patients were anomic following head injury and exhibited high levels of item consistency. Neither presented with any evidence for a comprehension deficit, although EE was not given the range of tasks used here. Their major naming error type was a failure to respond with occasional semantic errors and circumlocutions but no phonological errors. There are one or two differences between the two patients. EE was slightly more consistent than JS (EE inconsistently named 21/100 pictures, JS was inconsistent on 32/100). Although the improvement in JS's naming performance from single phonemic cues did not reach significance, it seems likely that they were relatively more effective for JS than for EE (who only named 4/40 extra items with a phonemic cue). Of course, it would be fascinating to know whether the multiphonemic cueing, or priming from a written word, used in this study would have proved as effective for EE as it did for JS.

The graded hypothesis provides a single explanation for many of the effects reported for JS. If the naming system is graded according to familiarity, frequency, or some combination of variables, then it should be possible to predict JS's consistency distribution. The distributions indicated by the linear regressions for both the Hundred Picture Naming Test and the Boston Naming Test were far from that observed. Although this could be used to argue against the graded hypothesis and for the loss of specific lexical entries, such a conclusion relies on the assumption that the relationship between accuracy and psycholinguistic factors is linear. It seems equally likely that the system (and the underlying neural structures) will behave in a nonlinear manner and consequently, the consistency distribution might be predicted by a nonlinear combination of the psycholinguistic variables.⁵

⁵ It was shown in the section entitled 'Naming Performance on the Hundred Picture Naming Test' that a stochastic model, which assumes that a proportion of items have been permanently lost, can be fitted to JS's consistency distribution. It is also possible to fit this pattern with a variable such as familiarity, for example, if it is first submitted to a sigmoid transform:

Throughout this paper distributed and localist models have been contrasted in a strict way, i.e., the localist models have been taken literally to predict the loss of specific lexical items (cf. Howard, 1995; Page, submitted). Although JS's initial data seemed to support this hypothesis, it is clear that the consistently unavailable items were not lost but were hard for him to retrieve. It should be noted that although some authors use localist representations within their architectures, it is possible that this kind of representation is adopted by most as a convenience approximation to some form of distributed representations contained in the brain. For example, Dell and his colleagues relate the localist lexical layer in their naming model to the set of hidden units between semantics and phonology in a distributed naming architecture (see Dell et al., 1997). Thus, in most approaches, the localist architecture would not allow for specific lexical items to be lost after brain damage (see footnote 1). Rather, an injury would lead to a more global type of impairment at a specific level of the model; for example, raising the threshold of each lexical item. If this interpretation is correct, then most of the data shown by JS can be accommodated within a localist model. One selection of typically unfamiliar items (with low resting levels) will be extremely hard to retrieve, i.e., consistently unavailable, while a group of slightly more familiar words may occasionally be able to rise above (the abnormally high) threshold at least on some occasions (the inconsistently named items) and the most familiar will be relatively robust to the effects of the brain damage so that they are nearly always activated (the consistently available words). The results of cueing and priming would follow from this revised localist theory; the two facilitating methods would result in a temporary rise in the activation of the target word making retrieval more likely, even for the consistently unnamed items. In summary, although localist and distributed models have different structures and functions, both can produce very similar behaviors after damage if they are treated as graded systems.

$$p(\text{retrieval}) = \frac{1}{[1 + e^{-(\text{familiarity}-b) \times c}]}$$

where b is the mean value (546) and c controls the angle of slope (set to 0.05). The original familiarity variable correlated ($r = 0.29$) with JS's naming accuracy, and the transformed familiarity variable correlated to a similar level ($r = 0.25$). When this variable is used to provide a probability for naming each item on 0, 1, 2 or 3 occasions (see the section entitled "Naming Performance . . ." for the precise method) the resulting predicted distribution [$E(0/3) = 32.1$, $E(1/3) = 15.0$, $E(2/3) = 17.4$, $E(3/3) = 35.5$] is very close to JS's observed distribution [$O(0/3) = 33$, $O(1/3) = 13$, $O(2/3) = 19$, $O(3/3) = 35$]; in fact, the two are not significantly different ($\chi^2 = 0.45$, $p = .93$). This is not to suggest that accuracy in the naming system is definitely sigmoid in character, but rather that it is possible to fit the consistency distribution with psycholinguistic variables if they are combined in a nonlinear fashion. It is entirely possible that other nonlinear functions may provide a better description of JS's data.

Are there any differences between the distributed and the revised localist architectures? It was argued above (see “Do Unnamed Items Predict Naming Failure on Other Pictures?”) that if speech production contains distributed representations then it should be possible to predict naming success for a group of items by the accuracy observed on a related set of pictures, although the exact prediction will depend on the locus of the specific patient’s impairment. Assuming JS had an impairment in the link between (intact) semantic representations and (intact) phonological representations, the highest degree of inter-item predictability should occur between semantically *and* phonologically related items, because these words have the greatest overlap in distributed representations at both relevant levels. JS failed to show a significant association between performance on pictures with only semantically or phonologically similar names. For those items that were both semantically and phonologically related, however, this predictability emerged. It is not clear that this result would emerge from the localist architectures. In their strict form, the loss of specific lexical items would not prevent the remaining lexical items reaching threshold. For example, if the word *cat* was deleted, although the related item *rat* would no longer benefit from the semantic and phonological feedback previously provided by *cat*, it would still reach threshold. Likewise, the “graded” localist alternative does not predict an association in naming accuracy between semantically and phonologically related pictures unless these items tend to be stored close together within a lexical “space”; thus, damage to a part of this lexical space will tend to effect similar types of word. This explanation is, however, a way to provide a localist approximation to an effect predicted by distributed models of speech production.

REFERENCES

- Allport, D. A. 1985. Distributed memory, modular subsystems and dysphasia. In S. K. Newman & R. Epstein (Eds.), *Current perspectives in dysphasia*. Edinburgh: Churchill-Livingstone.
- Butterworth, B. L. 1989. Lexical access in speech production. In W. Marslen-Wilson (Ed.), *Lexical representation and process*. Cambridge, MA: MIT Press.
- Caramazza, A., & Hillis, A. E. 1990. Where do semantic errors come from? *Cortex*, **26**, 95–122.
- Centre for Lexical Information 1993. *The Celex Lexical Database*. Nijmegen.
- Dell, G. S. 1986. A spreading activation theory of retrieval in sentence production. *Psychological Review*, **93**, 283–321.
- Dell, G. S., & O’Seaghdha, P. G. 1992. Stages of lexical access in language production. *Cognition*, **42**, 287–314.
- Dell, G. S., Schwartz, M. F., Martin, N., Saffran, E. M., & Gagnon, D. A. 1997. Lexical access in aphasic and nonaphasic speakers. *Psychological Review*, **104**, 801–838.
- Dunn, A. 1965. *The Peabody Picture Vocabulary Test*, Minneapolis, MN: American Guidance Service.

- Faglioni, P., & Botti, C. 1993. How to differentiate retrieval from storage deficit: a stochastic approach to semantic memory modelling. *Cortex*, **29**, 501–518.
- Franklin, S., Turner, J., & Ellis, A. W. 1992. *The ADA comprehension battery*. London: Action for Dysphasic Adults.
- Graham, K. S., Patterson, K., & Hodges, J. R. 1995. Progressive pure anomia: Insufficient activation of phonology by meaning. *Neurocase*, **1**, 25–38.
- Howard, D. 1995. Lexical Anomia: Or the case of the missing lexical entries. *The Quarterly Journal of Experimental Psychology*, **48(A)**, 999–1023.
- Howard, D., & Franklin, S. E. 1987. Three ways for understanding written words, and their use in two contrasting cases of surface dyslexia. In D. A. Allport, D. MacKay, W. Prinz, & E. Scheerer (Eds.), *Language perception and production: Common processes in listening, speaking, reading and writing*. London: Academic Press.
- Howard, D., & Franklin, S. E. 1988. *Missing the meaning?* Cambridge, MA: MIT Press.
- Howard, D., & Patterson, K. E. 1992. *The Pyramids and Palm Trees Test*. Bury St. Edmunds: Thames Valley Test Company.
- Kaplan, E., Goodglass, H., & Weintraub, S. 1976. *The Boston Naming Test*. Boston: Veteran's Administration.
- Kay, J., & Ellis, A. W. 1987. A cognitive neuropsychological case study of anomia. *Brain*, **110**, 613–629.
- Kucera, H., & Francis, W. N. 1967. *A computational analysis of present day American English*. Providence, RI: Brown University Press.
- Lambon Ralph, M. A., Sage, K., & Roberts, J. (submitted). Classical anomia: A neuropsychological perspective on speech production. *Cognitive Neuropsychology*.
- Levelt, W. J. M. 1989. *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
- Levelt, W. J. M. 1992. Accessing words in speech production: Stages, processes and representations. *Cognition*, **42**, 1–22.
- McClelland, J. L., & Rumelhart, D. E. 1981. An interactive activation model of context effect in letter perception: Part 1. An account of the basic findings. *Psychological Review*, **88**, 170–172.
- McClelland, J. L., Rumelhart, D. E., & Hinton, E. 1986. The appeal of parallel distributed processing. In D. E. Rumelhart, J. L. McClelland, & the PDP Research Group (Eds.), *Parallel distributed processing: Microstructure of cognition*. Cambridge, MA: MIT Press.
- Morton, J. 1985. Naming. In S. K. Newman & R. Epstein (Eds.), *Current perspectives in dysphasia*. Edinburgh: Churchill–Livingstone.
- Page, M. P. A. (submitted). Connectionist modelling in psychology: a localist manifesto. *Behavioural and Brain Sciences*.
- Plaut, D. C., Seidenberg, M. S., McClelland, J. L., & Patterson, K. 1996. Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review*, **103**, 56–115.
- Plaut, D. C., & Shallice, T. 1993. Deep dyslexia: A case study of connectionist neuropsychology. *Cognitive Neuropsychology*, **10**, 377–500.
- Quinlan, P. T. 1992. *The MRC Psycholinguistic Database*. London: Oxford University Press.
- Seidenberg, M. S., & McClelland, J. L. 1989. A distributed, developmental model of visual word recognition and naming. *Psychological Review*, **96**, 523–568.
- Snodgrass, J. G., & Vanderwart, M. 1980. A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, **6**, 174–215.

- Tyler, L. K. 1992. *Spoken language comprehension: An experimental approach to disordered and normal processing*. London: MIT Press.
- Warrington, E. K., & McKenna, P. 1983. *The Graded Naming Test*. London: NFER.
- Wingfield, A., Goodglass, H., & Smith, K. L. 1990. Effects of word-onset cueing on picture naming in aphasia: A reconsideration. *Brain and Language*, **39**, 373–390.